

Natural genetic variation of basal disease resistance in *Arabidopsis*

Andrew Glyn Staphnill

Disease and Stress Biology Department

John Innes Centre

Norwich

NR4 7UH

United Kingdom

A thesis submitted to the University of East Anglia
for the degree of Doctor of Philosophy

December 2009

© This copy of the thesis has been supplied on the condition that anyone who consults it is understood to recognise that its copyright rests with the author and that no quotation from the thesis, nor any information derived there from, may be published without the author's prior written consent.

Acknowledgements

It has been a privilege to have been given the opportunity to study at the world renowned John Innes Centre, which has been an exciting and stimulating environment to work in. During my studies I have relied on the kindness of many people for support and advice. I would firstly like to take this opportunity to thank my supervisor Chris Lamb, the Director of the JIC, who died suddenly as I was writing up this thesis and unfortunately never got to see it completed. Chris was a wonderful motivator as he empowered you to achieve even through difficulties and set backs with kind words and inspiring ideas. I would like to thank him for always believing in me and sharing with all of us his love of science and discovery.

I would like to thank my advisory team Jun Fan and Giles Oldroyd who have always given their time, tuition and advice freely and Gary Creissen, our lab manager, who has been on hand to teach and instruct me. Also, a big thanks to all the past and present members of the Lamb group, especially Ellen Colebrook and Matt Chabannes for making the lab such an enjoyable place to work.

All the members of the DSB department have been very helpful in offering feedback and comments during lab meetings and seminars and Luzie Wingen in particular has been a great help with some of the more complex statistics. I would also like to thank Volker Lipka and his group as well as Cyril Zipfel from the Sainsbury Lab for collaborative work and freely exchanging ideas. Colin Morgan and Ian Bancroft helped considerably in the early stages of the project in providing RIL lines and other seeds as well as an array of useful information. Jan Chojecki, the director of PBL, has been my CASE sponsor throughout this project and I would like to thank him for some very useful meetings and thought provoking suggestions. A special thanks to Jenni Rant for being a fantastic friend and for involving me with the Science, Art and Writing (SAW) teaching projects overseen by Anne Osbourn.

I have spent 8 wonderful years in Norwich completing my degree hat-trick and over all that time I have made a great number of friends who have encouraged me and made my time at university very enjoyable. There are far too many to name but I would particularly like to thank all my friends in the Athletics and Lacrosse clubs for all the laughs and memories (which alcohol hasn't obscured). I would like to thank my friend, Laura, for her great sense of humour and company during my PhD which helped to keep everything in perspective. All my housemates and people I've met through the Resident Tutor program over the years who have made the university campus a fun place to live and study. I would also like to especially thank Leila, whose love, support and understanding throughout my time at university has meant the world to me.

Finally, I would like to say the biggest thank you to all my family for their love and support. My mum, Christine, is the most devoted and selfless person that has been in my life and offers unending encouragement and belief that I can achieve my aims. My younger sister Hannah, who also studied in Norwich, is one of my best friends and a great musician to jam with. My thanks also go to my older sister, Joanne and her husband Jock, who have been an amazing help in hunting down those missing commas during my thesis corrections. I would also like to thank my Grandpa, Roy, who has been a massive part of my life and has motivated and encouraged me, often over a pint. To all these people and those I haven't had space to mention I would like to thank you for making all my years of study possible, I could not have done it without your good humour, help and support.

Abstract

Basal resistance utilises receptors which detect non-race specific elicitors known as pathogen associated molecular patterns (PAMPs). While not commonly resulting in a hypersensitive response, it shares some downstream signalling and effector components with *R* gene-mediated resistance. To better understand the genetic basis of basal resistance, a virulent *Pseudomonas* strain tagged with a constitutively expressed *LuxCDABE* operon was utilised to create a high throughput quantitative bacterial growth assay. Measuring bacterial photon emission from leaf tissue in a selection of ecotypes showed a wide range of bacterial growth, implying substantial natural genetic variation in basal resistance.

Several independent RIL populations were trialled and combined with other data for QTL analysis to define the chromosomal locations responsible for resistance variation. Within markers flanking each significant QTL peak, there are numerous candidate genes. Cross-referencing candidates against annotation information indicating involvement in biotic stress, respiratory burst or induction by flg22 treatment, creates a significantly refined list. A LysM receptor-like kinase family appears in the majority of QTL locations across independent populations at a statistically higher rate and is therefore implicated in pathogen recognition. Replication using other *Pseudomonas* bacterial strains produces similar LysM gene associations, implicating broad perception mechanisms.

Transcriptional, phylogenetic and functional characterisation found LysM genes consistent with characteristics of known disease resistance components. A LysM gene sequence comparison between genotypes with substantial variation of basal resistance found significant polymorphisms, however, a direct relationship between these differences and trait data could not be conclusively shown. *Pseudomonas AvrPtoB* kinase suppression can increase susceptibility in different genotypes but does not directly correlate with QTLs, LysM gene placement and suppression pattern.

Several T-DNA knockout lines showed a moderate increase (40-60%) in susceptibility for lines targeting specific LysM genes compared with Col-0 controls. Not all lines show this pattern but most tested lines show effects on basal resistance after gene knockout, directly implicating several LysM genes as components in basal resistance.

Table of contents

Acknowledgements	2
Abstract.....	3
List of figures and tables	7
Abbreviations.....	9
Chapter 1: General introduction.....	11
1.1 Plants and their environment.....	11
1.2 Plant defence	12
1.2.1 Physical and preformed defences	13
1.2.2 R gene mediated resistance	14
1.2.3 Defence responses to infection	15
1.2.4 Hormonal influence on pathogen defence	17
1.3 Basal resistance definition	20
1.3.1 Similarity between resistance pathways.....	24
1.3.2 PAMP interactions with PRRs.....	27
1.3.3 Hormonal signalling in basal resistance.....	29
1.3.4 Basal resistance suppression.....	30
1.3.5 Zigzag model	30
1.4 Natural variation.....	32
1.5 Economic importance of basal resistance.....	33
1.6 Experimental system.....	35
1.6.1 <i>Arabidopsis thaliana</i>	35
1.6.2 <i>Pseudomonas syringae</i>	36
1.6.3 Luciferase gene cassette.....	37
1.6.4 Recombinant inbred lines.....	38
1.7 Aims and objectives.....	39
Chapter 2: Materials and methods.....	41
2.1 <i>Arabidopsis</i> materials and methods.....	41
2.1.1 <i>Arabidopsis thaliana</i> genotypes.....	41
2.1.2 Plant growth.....	41
2.1.3 Media composition.....	42
2.2 Pathogen materials and methods	42
2.2.1 <i>Pseudomonas syringae</i> storage and growth	42
2.2.2 <i>Pseudomonas</i> infection.....	42
2.2.3 Infiltration controls and standardisation calculations	43
2.2.4 Refinement of technique experiments.....	43
2.2.5 Parental RIL line screening and selection of appropriate RIL lines.....	44
2.2.6 Visualising photon emission from Lux tagged bacteria using NightOWL luminescence camera.....	45
2.3 QTL analysis	45
2.3.1 QTL analysis information.....	46
2.3.2 Linkage mapping methods.....	46
2.3.3 QTL analysis software and files.....	47
2.3.4 Kruskal-Wallis QTL analysis.....	47
2.3.5 Interval mapping QTL analysis	48
2.3.6 Permutation tests in QTL analysis.....	48
2.3.7 Restricted MQM mapping and MQM mapping QTL analysis.....	48
2.3.8 Automatic Co-factor Selection in QTL analysis.....	49
2.4 Candidate gene identification and refinement.....	49
2.4.1 Candidate and disease related gene lists.....	50
2.4.2 Refinement of candidate lists and categorisation of genes.....	50
2.4.3 Candidate gene distribution within QTL regions.....	50
2.5 LysM gene characterisation	51
2.5.1 Annotation information	51
2.5.2 Phylogenetic analysis.....	51
2.5.3 Expression profiles of LysM genes.....	51

2.5.4 Sequencing	52
2.6 T-DNA knockout line identification.....	53
2.6.1 T-DNA insertion lines	53
2.6.2 DNA extraction.....	53
2.6.3 Oligonucleotide primers.....	54
2.6.4 Polymerase Chain Reaction	55
2.6.5 Agarose gel electrophoresis	55
Chapter 3: Refinement and characterisation of the luminescence-based <i>Pseudomonas</i> growth assay using <i>Arabidopsis</i> ecotypes	56
3.1 Introduction	56
3.2 Results.....	57
3.2.1 Relationship between Lux emission and bacterial number.....	57
3.3 Visualising the infection process	58
3.4 Parental RIL line testing.....	61
3.4.1 Technique refinement.....	64
3.4.2 Bacterial time-course of infection.....	65
3.4.3 Bacterial inoculum concentration.....	65
3.4.4 Age of bacterial inoculum	67
3.4.5 Plant age	68
3.4.6 Time of day influence on bacterial growth	69
3.4.7 Infiltration method	70
3.4.8 Environmental factors.....	71
3.5 Discussion	73
3.5.1 Natural variation in basal resistance of an <i>Arabidopsis</i> population.....	73
3.5.2 Refinement of technique	73
3.5.3 Visualisation of infection.....	76
Chapter 4: QTL analysis of basal resistance with <i>Arabidopsis</i> recombinant inbred lines	79
4.1 Introduction	79
4.2 Results.....	81
4.2.1 Selection of RIL lines	81
4.3.2 Basal resistance variation in recombinant inbred lines	81
4.3.3 Trait data from Sorbo x Gy-0 RIL population	82
4.3.4 Trait data from Nok3 x Ga0 RIL population	83
4.3.5 Trait data from Cvi0 x Ag0 RIL population.....	84
4.3.6 Trait data from Col-0 x Ler RIL population.....	85
4.3.7 Trait data from Wt-5 x Ct-1 RIL population.....	86
4.4 QTL analysis	87
4.4.1 Sorbo x Gy-0.....	88
4.4.2 Nok3 x Ga0.....	89
4.4.3 Cvi0 x Ag0.....	90
4.4.4 Wt-5 x Ct-1	91
4.4.5 Col-0 x Ler	92
4.4.6 Compiled QTL graphs.....	93
4.5 Selected components define QTL regions.....	93
4.6 Categorisation of refined candidate genes within the QTL regions.....	95
4.7 Correlation between QTL peaks and LysM gene location.....	97
4.8 LysM genes distribution statistics.....	98
4.9 Properties of the LysM gene family	100
4.10 Reanalysis of QTLs following Psm infiltration.....	100
4.11.1 Col-0 x Ler Psm	101
a) Col-0 x Ler Psm	101
b) Col-0 x Ler Pst.....	101
4.11.2 Nok-3 x Ga-0 Psm.....	103
a) Nok-3 x Ga-0 Psm	103
4.12 AvrPtoB.....	104
4.13 Discussion	106
4.13.1 QTL Analysis	106

4.13.2 Candidate gene refinement.....	107
4.13.3 Specificity of the basal resistance conferred by the QTLs.....	109
4.13.4 Role of LysM gene family in basal disease resistance.....	110
4.13.5 Effect of AvrPtoB on bacterial growth.....	111
Chapter 5: LysM gene family characterisation.....	114
5.1 Introduction.....	114
5.2 Assessing LysM gene function in basal resistance.....	117
5.3 Results.....	118
5.3.1 Summary of LysM genes and annotation information.....	118
5.3.2 BLAST searches of LysM genes.....	119
5.3.3 LysM gene phylogenetic trees.....	120
5.4 Expression profiles of LysM gene family.....	121
5.4.1 Developmental and anatomical profiling.....	121
5.4.2 Stimuli induced expression profiling.....	124
5.5 LysM gene sequencing.....	126
5.6 T-DNA insertion knockouts.....	127
5.6.1 Trait assessment of T-DNA lines.....	127
5.6.2 PCR assessment of T-DNA insertion type.....	127
5.6.3 Combined trait data and PCR insertion assessment.....	129
5.7 Discussion.....	132
5.7.1 Phylogenetic analysis and LysM gene properties.....	132
5.7.2 Transcriptional analysis.....	133
5.7.3 Sequencing analysis.....	134
5.7.4 T-DNA knockout lines.....	135
5.7.5 LysM gene effect.....	138
5.7.6 Hypothesis of LysM gene family involvement in basal resistance.....	139
Chapter 6: Discussion.....	141
6.1 Research summary.....	141
6.2 LysM gene family and PAMP recognition.....	145
6.3 Progression of the project and subject area.....	149
6.4 Application of the findings and commercialisation potential.....	150
6.5 Future work.....	151
References.....	156
Appendices.....	169
1.1 PCR primers.....	169

List of figures and tables

Table 1.1	15
Figure 1.1.....	23
Figure 1.2.....	27
Figure 1.3.....	28
Figure 1.4.....	31
Figure 1.5.....	37
Figure 1.6.....	38
Figure 3.1	58
Figure 3.2.....	59
Figure 3.3.....	60
Figure 3.4.....	61
Figure 3.5.....	63
Table 3.1	64
Figure 3.6.....	65
Figure 3.7.....	66
Figure 3.8.....	67
Figure 3.9.....	68
Figure 3.10.....	69
Figure 3.11.....	70
Figure 3.12.....	71
Figure 3.13.....	72
Figure 4.1.....	81
Figure 4.2.....	82
Figure 4.3.....	83
Figure 4.4.....	84
Figure 4.5.....	85
Figure 4.6.....	86
Table 4.1	87
Figure 4.7.....	88
Table 4.2	88
Figure 4.8.....	89
Table 4.3	89
Figure 4.9.....	90
Table 4.4	90
Figure 4.10.....	91
Table 4.5	91
Figure 4.11	92
Table 4.6	92
Figure 4.12.....	93
Table 4.7	95
Figure 4.13.....	96
Figure 4.14.....	97
Figure 4.15.....	98
Table 4.8	99
Figure 4.16.....	101
Table 4.9	102
Figure 4.17.....	103
Table 4.10	104

Figure 4.18.....	105
Figure 5.1.....	116
Figure 5.2.....	119
Figure 5.3.....	121
Figure 5.4.....	122
Figure 5.5.....	123
Figure 5.6.....	125
Figure 5.7.....	128
Figure 5.8.....	128
Figure 5.9.....	130
Table 5.1	131
Figure 6.1	149

Abbreviations

A, T, G, C	Adenine, thymine, guanine, cytosine
ACS	Automatic Co-factor Selection
ANOVA	Analysis of variance
At	<i>Arabidopsis thaliana</i>
Avr	Avirulence
bp	Base pair
°C	Degree Celsius
CC	Coiled coil
cDNA	Complementary DNA
Cfu	Colony forming unit
Chrom.	Chromosome
cm	Centimetre
cM	CentiMorgan
Col	Columbia
CPS	Counts per second
Df	Degrees of freedom
DNA	Deoxyribonucleic Acid
ET	Ethylene
F pr.	Significance of deviance ratio (F-test)
F1, 2	Filial 1, 2, ...
g	Gram
HR	Hypersensitive response
hrs	Hours
IM	Interval mapping
ISR	Induced systemic resistance
JA	Jasmonic acid
kbp	Kilo base pairs
l	Litre
Ler	Landsberg erecta
LLR	Leucine-rich repeat
LOD	Logarithm of odds
Log	Logarithm
Lux	Luminescence
LuxCDABE	Luciferase gene cassette
LysM	Lysine motif
M	Molar
MAP(K)	Mitogen-activated protein (kinase)
MAMP	Microbe associated molecular pattern
Mbp	Mega base pairs
MeJA	Methyl Jasmonate
µg, µl, µM	Micro gram, litre, Molar
mg, ml, mM	Milli gram, litre, Molar

MQM	Multiple-QTL models
m.s	Mean square
MS	Microsoft
MS	Murashige and Skoog media
NBS	Nucleotide binding site
RNS	Reactive nitrogen species
OD	Optical density
ORF	Open reading frame
PAMP	Pathogen-associated molecular pattern
PCR	Polymerase chain reaction
PR	Pathogenesis related
PRR	Pathogen recognition receptors
<i>Psm</i>	<i>Pseudomonas syringae</i> pathovar <i>maculicola</i>
<i>Pst</i>	<i>Pseudomonas syringae</i> pathovar <i>tomato</i>
PTI	PAMP induced immunity
qPCR	Quantitative polymerase chain reaction
QTL	Quantitative trait locus
R	Resistance
RAM	Random access memory
RIL	Recombinant inbred line
RLK	Receptor-like kinase
RLU	Relative light units
RNA	Ribonucleic acid
ROS	Reactive oxygen species
rpm	Revolutions per minute
S	Susceptibility
SA	Salicylic acid
SAR	Systemic acquired resistance
Sig.	Significance value – *0.1 **0.05 ***0.01 *****0.005 *****0.001
TLR	Toll-like receptor
UV	Ultra-violet
V	Volts
v.r	Variance ratio

Chapter 1: General introduction

1.1 Plants and their environment

Plants exist in environments surrounded by potentially pathogenic organisms, pressuring the evolution of plant defence responses that prevent successful infection for the majority of pathogens. In competitive environments, plants and pathogens have entered into an evolutionary arms race which has forced plant production of a diverse array of pathogen receptors and multiple defence mechanisms (Hammond-Kosack and Jones, 1997). Consequently, pathogens have evolved complex pathogenesis effectors to overcome plant defences (Shan et al., 2007). The result of this competition between the plant and the pathogen is a highly specialised environment in which usually only species specific interactions can occur and successful plant infection is relatively rare. Some non-host interactions do still occur however infection tends to be opportunistic and relies on wounding or insect vector transfer (Hammond-Kosack and Jones, 1997).

Modern civilisation and agriculture rely on growing crops at high densities on relatively small areas of farm land, using intensive farming methods to increase yields and final product output (Labanna and Banga, 1993). Advances have been made in breeding crop cultivars that are high-yielding, hardy and require minimal refinement to create a final product (Labanna and Banga, 1993; McDonald and Linde, 2002). Despite these innovations and the advent of a wide variety of chemical herbicides, pesticides, fungicides and anti-bacterial treatments, much crop yield is lost as a result of plant disease. Demand for effective disease control led to the development of crop varieties that contained single major resistance genes (*R* genes) that block infection from certain races of pathogens (McDonald and Linde, 2002). *R* gene-mediated defence displays strikingly resistant phenotypes, which have previously been well-characterised and introduced commercially; although initially very successful, they proved ineffective in field environments over time (McDonald and Linde, 2002). Breeders and farmers require new sources of effective, wide-ranging and durable forms of disease resistance to maintain and increase yield outputs.

Basal disease resistance plays an important role in the plant defence network and offers a non-race-specific response of varying degrees of intensity to invading pathogens. While basal resistance is significantly weaker than the single major *R*

gene defence, in that it does not typically activate programmed cell death, it can however slow and reduce pathogen progression and remain durable over generations (Katagiri et al., 2002). Basal resistance signalling pathways and pathogen recognition mechanisms are not well understood and are under-represented in current literature, with only a few well-defined examples (Zipfel et al., 2006; Chinchilla et al., 2007). The range of variation in resistance due to basal mechanisms has not been well-characterised however, existing data does suggest a polygenic system with multiple receptors and pathways. Location and characterisation of critical components in basal pathways are important prerequisites before these mechanisms can be enhanced to confer resistance in commercially-grown varieties. These topics will be addressed in this thesis.

1.2 Plant defence

Plant defence involves a network of complex, tightly-regulated pathways that often overlap and influence each other, and are completely integrated into plants' developmental and lifecycle strategy (Katagiri et al., 2002). Defence begins at the surface of a plant. Physical adaptations limit the access of potential pathogens and deter herbivore and insect feeding, which are common entry points for pathogens (Thaler et al., 2002).

Pathogens must then contend with a variety of metabolic defences and detection mechanisms that can differentially activate defence compounds and proteins, depending on the invading pathogen (Pieterse and van Loon, 1999). There are several overlapping metabolic defences described in the literature, typically with distinctive characteristics at some point along their pathways to distinguish their pathway or mode of action. However, distinctions between pathways are hampered by conflicting experimental evidence (Dangl and Jones, 2001). Therefore, this thesis will firstly examine the whole sphere of plant defence in order to define and place basal disease resistance into context within other plant defence networks.

1.2.1 Physical and preformed defences

Physical plant defences such as tree bark, spines and trichomes provide general defence against herbivores and herbivorous insects by deterring or slowing feeding (Taiz and Zeiger, 2002). Preformed toxic compounds such as terpenes, alkaloids and glycosides act as deterrents to herbivorous feeding and can reduce physical damage in addition to reducing a plant's contact with potential pathogenic vectors. Defensive preformed barriers, such as a leaf's thick waxy acyl lipid cuticle, stomatal morphology and the cellulose-rich cell wall, also reduce pathogen contact with viable host cells (Taiz and Zeiger, 2002). Physical and physiological adaptations can substantially influence pathogenic susceptibility in a natural environment (Glazebrook, 2001).

Preformed secondary metabolite compounds with antimicrobial or inhibitory properties such as phytoalexins, phytoanticipins, saponins and glucosinolates form a preliminary line of plant defence. Defensive compounds are not evenly dispersed as distribution and composition are often tissue specific e.g. relatively higher storage concentration in trichomes or epidermal layers (Osbourn, 1996, 1999). Treatment with bacterial flagellin has been associated with an accumulation of glucosinolates products around the infection site and in *Arabidopsis*, the secondary metabolite camalexin has been shown to contribute to disease resistance through disruption of bacterial membranes (Bednarek and Osbourn, 2009). These antimicrobial secondary metabolites are often present in all tissues at background levels. However in response to microbial attack, can be locally synthesized to limit disease development through non-specific inhibition of protein synthesis or other mechanisms (Bednarek and Osbourn, 2009). Pathogenic adaptation can overcome preformed defence when, for example, saponin-detoxifying enzyme hydrolyses saponin and mediates defence suppression (Bouarab et al., 2002).

1.2.2 *R* gene mediated resistance

Successful pathogen colonisation of plant tissue is referred to as a compatible interaction with a virulent pathogen. An incompatible interaction indicates that the pathogen is avirulent and perception of its presence has stimulated a defence reaction in the host plant (Katagiri et al., 2002). Determination of a compatible plant / pathogen interaction follows the Mendelian “gene-for-gene” model, in which resistance occurs only when a dominant monogenic host resistance gene (*R* gene) interacts with a corresponding pathogen avirulence gene (*Avr* gene) (Flor, 1955; Nimchuk et al., 2001). In a compatible interaction, pathogenic *Avr* genes facilitate growth and development within the host tissue and are often essential for successful infection, therefore, *Avr* genes are often highly conserved in the pathogen population (Veronese et al., 2003). Plants have evolved membrane-bound and cytosolic receptors which recognise the conserved protein regions or ligands, known as elicitors, in a variety of fungi, bacteria and viruses (Dangl and Jones, 2001).

Gene-for-gene interactions typically involve a peripheral membrane protein receptor that spans or resides on the cytoplasmic face of the plasma membrane e.g. RPM1, RPS2, RPP1 (Boyce et al., 1998). Such receptors predominantly consist of a Leucine rich repeat (LLR) sequence, a nucleotide binding site (NBS) and either a coiled coil (CC) pattern or a motif conserved with Toll-like receptors (TLR) (Meyers et al., 1999; Ellis et al., 2000; Belkhadir et al., 2004). These receptors are encoded by families of *R* genes, such as the resistance against *Peronospora parasitica* (*RPP*), resistance to *Erysiphe cichoraceum* (*RPW*) or resistance against *Pseudomonas syringae* (*RPS*) groups, which often share conserved characteristics (Aarts et al., 1998; Meyers et al., 1999). These characteristics include the specific recognition interaction of the corresponding *Avr*-gene-related ligands, signalling cascade activation, and the capacity for rapid evolution of specificity (Chelkowski and Koczyk, 2003). Host-specific *R* genes encoding NBS-LRR proteins recognise race-specific pathogenic *Avr* gene products, either directly or through intermediates, and confer the most effective defence reaction to prevent pathogen growth through a signal transduction cascade (Nimchuk et al., 2001; Katagiri et al., 2002). A direct physical interaction has been shown between: RRS1-R, a protein conferring resistance to bacterial wilt; PopP2, a type III effector and between the *Pi-ta* gene product, conferring resistance to strains of the rice blast fungus and the corresponding *Avr-Pi*-

ta gene product (Jia et al., 2000; Deslandes et al., 2003). A defence induced signalling cascade activates expression of defence proteins through one of three pathways: an *EDS1*, *PAD4*-dependent pathway, a *NDRI*, *PBS2*-dependent pathway or through a third pathway that is independent of the two downstream constraints previously mentioned (Feys et al., 2001; Tornero et al., 2002). *R* gene-mediated resistance is typically associated with a hyper-sensitive response (HR), whereby the host-coordinated cell death around the site of infection creates a lesion which prevents further pathogen proliferation through deprivation of resources (Heath, 2000).

	Plant host <i>R/R</i> or <i>R/r</i>	<i>r/r</i>
Pathogen <i>Avr</i>	Resistant	Susceptible
<i>avr</i>	Susceptible	Susceptible

Table 1.1 Mendelian gene-for-gene interaction table displaying corresponding *Avr* and *R* gene components for successful recognition and resistance response.

1.2.3 Defence responses to infection

Plants lack a centralised adaptive immune system or specialised cells for disease resistance, as is common in vertebrates, therefore each plant cell must be capable of producing or inducing a defence reaction (Veronese et al., 2003). Many defence compounds are damaging to plant tissue and consequently need to be tightly regulated in order to avoid interfering with plant development (Dangl and Jones, 2001). As pathogen populations are constantly evolving to overcome and adapt to plant defences, a multilayered network of inducible defence has formed to counteract this attack (Chisholm et al., 2006). Consequently, wide varieties of compounds with differing modes of action are readily expressed upon detection of an elicitor, however there are differences and similarities in the components, composition and strength of response, depending on which inducible defence network instigated the reaction.

There are substantial similarities between components of *R* gene dependent and basal resistance responses, however basal defence is typically of a less vigorous nature (Zipfel, 2008). Conserved components of both pathways include ionic flux

within the cell following pathogen perception, which is primarily a result of calcium fluctuation, and acts to coordinate initial signalling (Abramovitch et al., 2006a). Downstream of the initial defence compound release, a set of mitogen-activated protein (MAP) kinase genes are expressed, as is the localised synthesis of ethylene, which is involved in defence signalling and co-ordination (Xiao et al., 2005). Phytoalexins and phenolic molecules are released in close proximity to the plant cell wall to limit the action of pathogenic digestive enzymes around the infection site and *in vitro* studies have showed their antimicrobial properties on various bacterial species (Osbourn, 1999; Bednarek and Osbourn, 2009). Following infection, callose deposition occurs and lignification of the cell wall is stimulated, which fortifies the wall from digestive enzyme exudates secreted by the pathogen in both *R* gene and basal mediated resistance (Tang et al., 1999).

In addition to the previously described components common to both *R* gene-mediated resistance and basal defence (e.g. phenolics), an oxidative burst at the site of infection results in the accumulation of reactive oxygen species (ROS). These small, highly reactive molecules include the superoxide anion O_2^- , H_2O_2 , OH^- and $^1O_2^-$ (Lamb and Dixon, 1997; Rentel et al., 2004). ROS accumulation can cause damage to the pathogen through non-specific oxidation of lipid membranes and accessible protein side chains, in addition to enzyme disruption and DNA mutation (Lamb and Dixon, 1997; Thordal-Christensen et al., 1997; Torres et al., 2002). Reactive nitrogen species (RNS) react and function similarly (VanderVliet et al., 1997). In addition to direct defensive action, these compounds also function as a substrate for oxidative cross-linking of structural proteins within the cell wall. This toughens this barrier to protect from pathogenic chemical degradation (Brisson et al., 1994).

The collective term pathogenesis-related (PR) protein describes a wide array of defensive enzymes and other proteins that are specifically induced during contact with a pathogen. Proteins synthesised are classified, based on function, into 17 categories which include glucanases and chitinases, proteinase-inhibitors and peroxidases, in addition to oxidases and antifungal compounds (Stintzi et al., 1993). This pathogen-induced group of proteins represents the largest change in protein composition after a pathogen is recognised. PR proteins work additively to produce a combined defence effect that attacks the pathogen cell wall, both actively, in the case of chitinases, and passively by slowing pathogen reactions, for example through inhibition of proteinase (Stintzi et al., 1993).

A significant difference between basal and *R* gene reactions is the hypersensitive response (HR). A HR encompasses the infection site and restricts nutrients to the developing pathogen which leaves a characteristically isolated area of necrosis surrounded by uninfected healthy tissue (Bittel and Robatzek, 2007). However in a compatible interaction, HR may still be induced, albeit too slowly to contain the pathogen spread. A slowly increasing HR lesion develops, tracking the disease progression (Morel and Dangl, 1997; Heath, 2000).

Assay systems designed to compare differences between basal and *R* gene resistance must take these similarities into account and tightly define the ranges of response for each defence mechanism. Comparative assays should also be quantitative and unambiguous so that differences in assay results are clearly attributable to defensive mechanistic action. Such constraints and solutions are discussed in more depth in subsequent assay system sections.

1.2.4 Hormonal influence on pathogen defence

Plant defences are tightly coordinated via gene expression, transcription factors and many other influential components which act to induce and optimise the defence response to a specific pathogen (Feys and Parker, 2000). Microarray expression-profiling and proteomic studies show that cooperative hormonal fluctuations and cross-talk between salicylic acid (SA), jasmonic acid (JA) and ethylene (ET) have a substantial influence on defence related gene expression and can modify the composition of a defence reaction in response to wounding, fungi, viruses or bacteria (Koornneef and Pieterse, 2008). For example, an interaction between PAMP and SA signalling is evident as both inoculation with *Pst* DC3000 and benzothiadiazole (a functional analogue of SA) result in the induction of FRK1, a flagellin-induced receptor kinase (Flors et al., 2008; Tsuda et al., 2008). Comparative microarray analysis of transcriptional changes in *npr1* mutant plants overexpressing an NPR1-GR (glucocorticoid receptor) fusion protein indicated that several WRKY genes are direct transcriptional targets of NPR1 (Eulgem, 2006; Wang et al., 2006). The introduction of SA and dexamethasone, a synthetic functional homologue, activated the NPR1 pathway and triggered translocation of NPR1-GR to the nuclei however, by inhibiting protein biosynthesis with cycloheximide transcription of indirect NPR1 targets. T-

DNA knockout mutants of the candidate WRKY components were then used to confirm their roles in NPR1 signalling (Eulgem, 2006; Wang et al., 2006). Microarray analysis of changes in transcription of 2,375 defence associated genes following inoculation with *Alternaria brassicicola* revealed that 50% of the genes induced by ET are also induced by JA, 55 genes were co-induced by SA and JA and 126 genes were induced when exposed to either SA, JA or ET which suggests overlapping hormonal influence (Schenk et al., 2000). The antagonistic relationship between hormones is demonstrated by 8 genes, including a chalcone synthase and a lipid-transfer gene, which were significantly induced by SA treatment but significantly repressed by a JA orthologue (Schenk et al., 2000).

SA is locally induced in an incompatible response to pathogen infection primarily by *SID2* gene which encodes an isochorismate synthase (Wildermuth et al., 2001). The resulting SA accumulation around the site of an infection acts to regulate expression of defence compounds, typically via TGA and WRKY transcription factors (Tsuda et al., 2008). High relative SA concentrations around the infection site prevent senescence through inhibition of ET biosynthesis and also act antagonistically to initially block wound responses, presumably to prevent synthesis of wound response proteins when pathogen-related protein synthesis is required (Van Wees et al., 2000; Huang et al., 2005; Tsuda et al., 2008). SA levels fall rapidly post-infection, which releases an inhibition of ET and JA-mediated wound responses (Van Wees et al., 2000). SA has also been implicated in the systemic acquired resistance (SAR) pathway that signals distal plant tissues to prepare defences for potential pathogenic attack (Ryals et al., 1996).

JA and methyl jasmonate (MeJA) are plant hormones derived from lipid biosynthesis and are noted to induce senescence at high concentrations, as well as influencing the development of flowers and fruit (Penninckx et al., 1998; Devoto and Turner, 2003). At low concentrations, for example, JA can alter protein and mRNA production of chalcone synthase, sugar storage enzymes and a proline rich cell wall structural repair protein (Creelman et al., 1992). JA has been determined as a critical component in many stress response pathways, particularly those associated with the mediation of plant wounding responses (Creelman et al., 1992). During wounding, JA levels rise significantly however, JA expression during pathogen attack is variable and possibly race-specific (Devoto and Turner, 2003). JA is capable of inhibiting SA signalling at high concentrations therefore cross-talk involving SA and JA is likely to

finely regulate and differentially activate plant defences (Penninckx et al., 1998; Devoto and Turner, 2003).

ET is biologically synthesised from methionine and is involved in fruit ripening, senescence of leaves and induction of root hair development. ET also acts as a plant hormone involved in plant defence and is recognised by cells by means of five homologous membrane-bound receptors that negatively regulate signal binding (Penninckx et al., 1998). In the presence of ET, a mitogen-activated protein (MAP) kinase is inactivated which initiates a signal transduction pathway involving various components such as MPK6 and EIN3, as well as transcription factors such as ERF1. The resulting signal stimulates chitinase and *PR* expression (Ton et al., 2001). ET action is consistently linked with JA. Although the interaction does not necessarily involve an up-regulation of JA, it is thought that slight increases of ET in the tissue increase cellular sensitivity to JA. This may explain JA and ET's mutual dependence in addition to the observation that JA and ET levels may only vary slightly during a defence reaction (Penninckx et al., 1998). JA and ET are commonly associated with regulation of wound response expression and can, at high concentrations during a large wounding response, suppress SA accumulation therefore prioritising wound response over pathogen defence (Creelman et al., 1992).

JA and ET act in unison during a wound response and have been known to stimulate SA production with corresponding up-regulation of *PR* gene defence, suggesting that after a wound defence is appropriately instigated, plants prepare the wound site for possible pathogen attacks (Feys and Parker, 2000). The majority of pathogens rely on opportunistic breaches in a plant's defence to infect and therefore, preparation of a wound site for pathogenic attack is a valid strategy.

Systemic defence responses also act antagonistically: in tobacco, JA-mediated wound responses are blocked through elevated SA concentrations (Dangl and Jones, 2001). This strategy has also been noted by Moran and Thompson (2001) in cucumber displaying fungal-induced systemic acquired resistance (SAR). In this instance, induced SAR is associated with enhanced susceptibility to herbivory by both beetles and aphids. Conversely, tobacco plants with reduced SA accumulation, caused by gene silencing, show reduced SAR and consequently herbivore resistance increases. This evidence implies that a plant that has suffered a primary pathogen infection typically prioritises SAR above small wound responses (Moran and Thompson, 2001).

A resistance response can be activated via two distinct signalling pathways that are commonly differentiated by their utilisation of salicylic acid (SA). The SA-dependent pathway has been studied using a *NahG* transgenic line, which is unable to accumulate SA. Wild-type plants typically activate a resistance response via the implicated regulators of SA-dependent defences; PAD4 and EDS1 (Hammond-Kosack and Jones, 1997; Zhou et al., 1998). The SA accumulation mutants are more susceptible to a test screen of pathogens and fail to produce a local pathogenic response or systemic signal to other cells. PR defence compound production is not always associated with SA production, and defence reactions are not always consistent with the reactions induced by SA (Fouts et al., 2003; Huang et al., 2005). Up-regulation of one or more of the *PR* genes (*PR-1*, *PR-2*, or *PR-5*) is used as a marker for defence activation of the SA-dependent pathway, however in *NahG* plants, unable to accumulate SA, induction of *PR* genes indicated that a parallel pathway exists which can induce defence reactions without SA.

The JA-insensitive mutant *jar1* and ET-insensitive mutants *etr1-1* demonstrated that SA-independent systemic resistance can be induced upon pathogen infection, is dependent on JA and ET and leads to the expression of the defence gene *PDF1.2* (Clarke et al., 2000). Resistance can be induced in *Arabidopsis* to *Alternaria brassicicola* and *Botrytis cinerea* with prior exposure to MeJA, whereas in contrast, *jar1lein2* mutants exhibit increased susceptibility to the same pathogens (Clarke et al., 2000). Both the SA-dependent pathway and JA-dependent pathways are capable of differentially activating a variety of defence compounds in order to optimise the defence reaction (Pieterse and van Loon, 1999). The full complexity of the genes involved, signal pathways and the interplay between them is not clearly understood. However, its major components appear to be active in both *R* gene and basal-induced immunity (Tsuda et al., 2008).

1.3 Basal resistance definition

Basal resistance is a collective term to describe a series of early plant recognition and defence response pathways following pathogen exposure. *R* gene interactions, typically characterised by a HR response, are not associated with a basal defence response (Zipfel et al., 2006). However, a genotypically variable physiological reaction is induced and acts to reduce pathogen proliferation. Basal resistance also is

referred to as including, but not wholly defined by, pathogen-associated molecular patterns (PAMP)-triggered immunity (PTI) as well as effector triggered immunity (ETI) (Jones and Dangl, 2006). There are significant gaps in our understanding of plants' multiple and continuous defensive layers when exposed to pathogens (Jones and Dangl, 2006). This makes precise definitions and physiological limits difficult to establish, however it is important to clarify some boundaries within the confines of this thesis. Basal disease resistance can be defined as a non-race-specific reaction to both virulent and avirulent pathogens that act to limit pathogen growth using a variety of broad-ranging defence products, but not typically the hypersensitive response (Zipfel et al., 2004). It may not prevent infection by a virulent pathogen, but it may limit disease development enough to allow the plant to complete its life cycle (Katagiri et al., 2002). Even with a gene-for-gene interaction, the absence of basal resistance can allow even an avirulent pathogen to create disease symptoms therefore, basal mechanisms are involved in the majority of pathogen interactions (Katagiri et al., 2002). Basal resistance offers a general and broad defence reaction that is not always associated with cell death, chlorosis or other visible physiological symptoms (Zipfel, 2009). A detectable molecular change in a wide range of defence compounds can be induced by this reaction to limit the current infection and aid prevention of subsequent infection, with the most prevalent being the induction of PR genes and proteins (Varga and Szegedi, 2007). Several papers have used physiological assays to characterise basal responses, such as a reduction in vascular flow rate, callose deposition and an oxidative burst (Rentel et al., 2004; Oh and Collmer, 2005; Asselbergh and Hofte, 2007). However, the biological roles of these responses in basal resistance are still not clear, therefore, unambiguous assays of pathogen growth are preferable.

Basal resistance is initiated following detection of the small, conserved molecules derived from pathogens. These molecules which are often fundamental to pathogenesis, contain invariable epitopes conserved across microbial species and taxa, such compounds are often referred to as pathogen-associated molecular patterns (PAMPs) (Chisholm et al., 2006; Bittel and Robatzek, 2007). Host plants are thought to recognise PAMPs, such as the gram-positive elicitor lipoteichoic acid and bacterial lipopolysaccharides, using membrane-bound receptors recognized by receptor-like kinases (RLKs), Toll-like receptors (TLRs) and other pattern recognition receptors (PRRs) (Espinosa and Alfano, 2004; Zipfel and Felix, 2005). For example, the

flagellin peptide elicitor flg22 is perceived by the leucine-rich repeat (LLR) kinase FLS2 and triggers basal defence reactions that show a strong similarity to *R* gene-mediated responses (Zipfel et al., 2004). Arabidopsis *FLS2* gene transfer and over expression in tomato was shown to provide flagellin perception in wild-type naive tomato plants where FLS2 and flg22 compounds were shown to co-precipitate *in vitro* which suggests a physical association (Bittel and Robatzek, 2007). FLS2 is typically localised to the external face of the plasma membrane and following flg22 exposure and receptor phosphorylation the FLS2 internalises in an endocytic process to predominantly the internal face of the membrane where further WRKY associated signalling occurs (Bittel and Robatzek, 2007). Other non-specific elicitors are implicated in activating basal defence such as host cell wall fragments, osmotic stress, alkalinisation of the cytoplasm, chitin, fungal ergosterol, peptidoglycan and components of the type III secretory pathway (Kim et al., 2005b; Oh and Collmer, 2005). Compounds commonly found in viruses, such as double-stranded RNA, are also associated with basal defence induction, which suggests that basal defence can recognise and defend against a broad range of pathogens without specific *R* gene-mediated interaction (Kim et al., 2005b; Oh and Collmer, 2005). Rapidity of defence activation is a useful distinction between mechanisms: basal defences have been noted to activate in less than 10 minutes after initial contact. Experimental introduction of flagellin, lipopolysaccharide or EF-Tu PAMPs to cell suspensions show increased extracellular pH as a result of ion flux, increased ROS and NOS, in addition to MAPKs activation and *FRK1*, *WRKY29* expression increases (Abramovitch et al., 2006a). In contrast *R* gene mediated response is typically a later but stronger reaction, induced in 2-3 hours following cytoplasmic detection of *Avr* gene products introduced via the type III secretory system (Abramovitch et al., 2006a). This has been shown experimentally, primarily through inoculation of *R*-gene-expressing leaves with bacteria that express a corresponding effector *Avr* gene. This results in a similar composition of defence compound induction as basal resistance with the notable addition of HR (Abramovitch *et al.*, 2006).

Virulent and avirulent pathogen-/host interactions may not necessarily be qualitatively different from the specific *R* gene-mediated resistance; therefore aspects of the basal mechanism may not be distinctly different from the incompatible interactions (Navarro et al., 2004). *R* gene receptors typically share common mechanistic action and molecular variation within an *R* gene class, therefore, a host

may have a similar *R* gene with low affinity for the *Avr* gene products (Meyers et al., 1999). Consequently, a weak interaction occurs between pathogen elicitors and host receptors which induces a weaker defence response that slows a compatible interaction, but does not prevent pathogen development through a HR (Heath, 2000). Specific categorisation and separation of defence pathways is not always possible given our current level of understanding, it is clear that many mechanisms and components of recognition and reaction are yet to be elucidated.

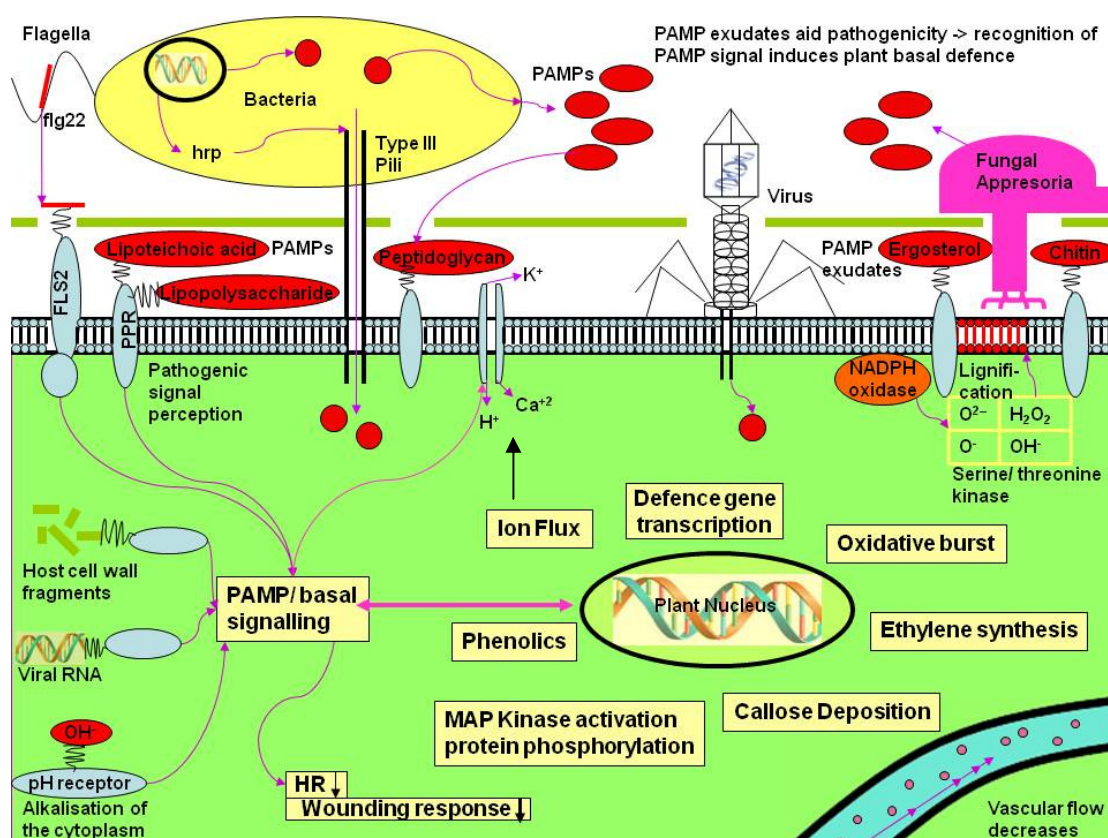


Figure 1.1 Simplified interpretation of plant, pathogen and PAMP interaction in basal resistance pathways and the subsequent defence reactions (similar to those of *R* gene) elicited by successful recognition and signalling. Diagram shows commonly described PAMPs, such as flg22 attaching to pathogen recognition receptors such as FLS2 leading through unknown signalling pathways to defence responses. Common PAMP-induced defences are shown, including oxidative burst, phenolics and callose deposition. Although these defences are similar to *R* gene defence components, the ranges and extents may differ in basal-induced resistance.

1.3.1 Similarity between resistance pathways

The genes involved in basal resistance are only beginning to emerge in the literature and pathways have yet to become clearly distinct. There are many downstream areas of similarity between layers of defence as basal resistance has also been shown to utilise components from SA-dependent pathways. Mutations in *PAD4*, *EDS1*, *EDS5*, *NPR1* and *SID2* have all shown a consequential reduction, or loss, of basal disease resistance when confronted with a pathogen (Feys et al., 2001). Basal resistance is also impaired when the *COI1* and *PDF1.2* genes' function is disrupted (Penninckx et al., 1998; Moran and Thompson, 2001). This suggests that basal resistance may be able to activate both SA-dependent and JA/ET-dependent pathways downstream of a possibly unique upstream recognition and signalling component which would differentiate basal resistance from other mechanisms (Penninckx et al., 1998; Moran and Thompson, 2001).

Following pathogen exposure, both *R* gene (e.g. *RPW8*) and basal resistance receptors, including FLS2, EFR and CERK1, appear to converge to exploit mitogen-activated protein kinases (MAPKs) and WRKY transcription factors in a signal cascade that coordinates, amongst other things, *PR* and *PRR* gene up-regulation (Fig.1.2) (Abramovitch et al., 2006a; Kwon, 2010). Both FLS2 and RPW8 perception of pathogens require the induction of critical components required for general resistance against bacterial pathogens such as *NHO1* and *FRK1* genes or the Flavin-dependent monooxygenase *FMO1* gene, which is an essential component of induced Systemic Acquired Resistance (Lu et al., 2001; Abramovitch et al., 2006a; Mishina and Zeier, 2006). Potential for substantial cross-talk between mechanisms is demonstrated with the example of the *RPW8* dominant *R* gene, which is commonly associated with mildew perception function and induction of HR (Xiao et al., 2001). This gene reportedly recruits components of basal defence in its induction of mildew defence, such as *EDR1* gene, which encodes a MAPKK kinase and appears to negatively regulate HR through suppression of *RPW8* transcription in early response to the pathogen (Fig.1.2) (Xiao et al., 2005).

It is likely that basal resistance is dependent on a wider array of receptors to detect a greater multitude of PAMP molecules than *R* gene pathways alone. *R* gene / *Avr* gene product interactions are highly specific so a large selection pressure has led to their evolution and maintains their presence in the plant population (Ellis et al.,

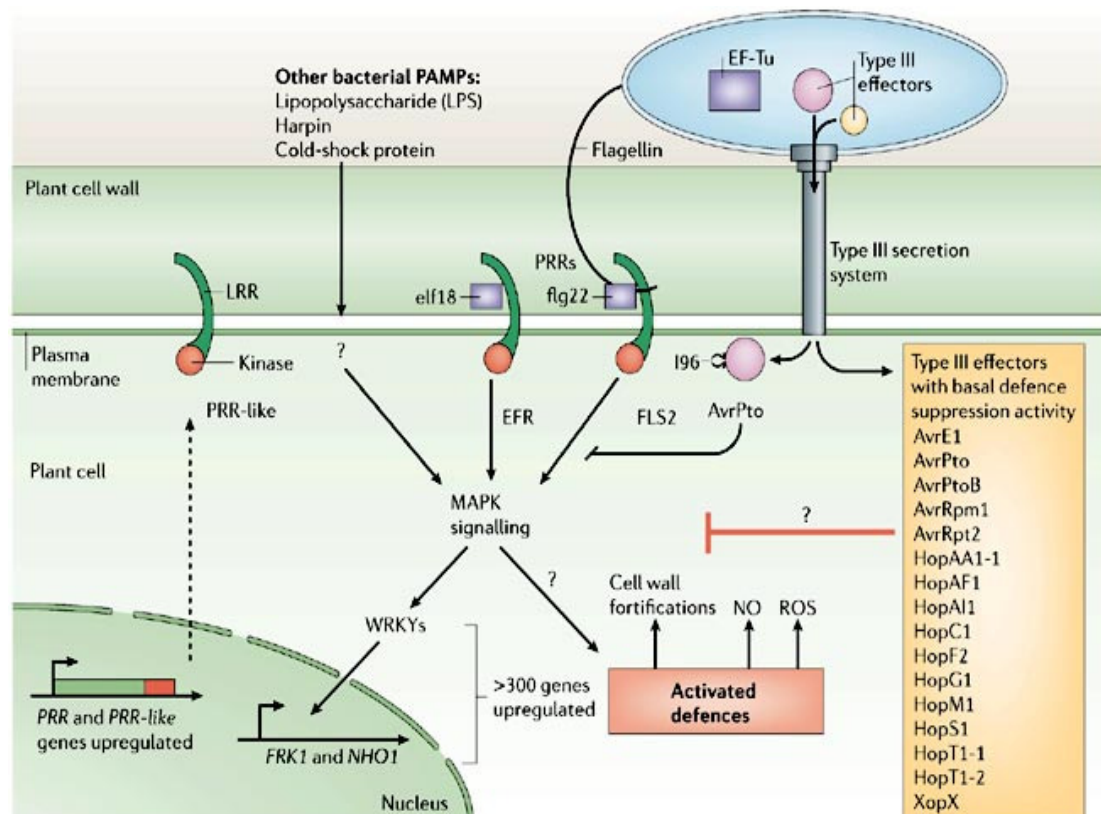
2000; Bittel and Robatzek, 2007). Basal resistance mechanisms require broad perception or a large range of receptors to perceive conserved but diverse pathogen elicitors. Such individual pathogen receptors do not have the selection pressure potential *R* genes have on pathogen populations, which would result in pathogens quickly overcoming their action. Adaptation over time has in likelihood led to high diversity of basal resistance molecular recognition components which benefit the host plant. However individually, high diversity has a relatively low impact on overall resistance when compared to *R* gene products (de Meaux and Mitchell-Olds, 2003).

The components of the basal resistance recognition of pathogens and associated defence reaction is only recently emerging in the literature, however, EF-Tu interaction with EFR is a particularly well known example in this area. EF-Tu is an abundant elongation factor found in all bacterial cells and acts as a PAMP which can be perceived by LRR-RLK EFR receptor in *Arabidopsis* (Fig. 1.2) (Kunze et al., 2004). The bacterial *EF-Tu* PAMP gene and elf18 peptide products contains a highly conserved 18 amino-acid sequence region from the N-terminus, which is critical to *EFR* perception, and can itself initiate a basal defence reaction (Chinchilla et al., 2007). *Arabidopsis efr* mutants display increased bacterial susceptibility as *Agrobacterium* transformation rates were significantly increased. Following detection of the elf18 peptide the EFR receptor activation leads to a signalling cascade via MAPK and WRKY genes which activate basal defence responses. (Kunze et al., 2004; Zipfel et al., 2006). EFR detection of EF-Tu via MAP kinase activation, initiates downstream defence gene induction, oxidative burst, callose deposition in addition to local SA synthesis and the inhibition of seedling growth (Nekrasov et al., 2009). EFR is associated with the plasma membrane but requires the stromal-derived factor-2 (SDF2) which associates with the endoplasmic reticulum to function in a protein complex to prevent the degradation of EFR (Nekrasov et al., 2009). A *dde2*, *ein2*, *pad4* and *sid2* quadruple mutant shows the flg22 or elf18 induced ETI signalling pathway is 80% dependent on these four genes (Tsuda et al., 2009). AvrRpt2 induced ETI also shows an approximate 80% dependency on these four genes, essential for JA and SA biosynthesis and the majority of ET responses which suggests a critical convergence point for both *R-gene* and PTI signal transduction to defence activation (Tsuda et al., 2009).

The *CERK1* gene found in *Arabidopsis* contains 3 LysM domains, RLK function and is predicted to be bound to the plasma membrane where it is involved

with the perception of a critical component of the fungal cell wall, chitin (Miya et al., 2007). Knockout mutants of *CERK1* gene show a complete loss of the ability to perceive the chitin PAMP and elicit a defence reaction, such as an oxidative burst or MAPK activation (Miya et al., 2007). CERK1 may also be involved in the perception of a critical component in the bacterial cell wall, peptidoglycan (Buist et al., 2008). Supporting the evidence that the CERK1 protein product is also involved in bacterial perception is that the type III effector protein AvrPtoB ubiquitinates the CERK1 kinase domain and so disrupts protein structure and function (Gimenez-Ibanez et al., 2009).

A similar gene found in Rice, encoding the chitin elicitor-binding protein (CEBiP), also contains LysM domains and is bound to the plasma membrane and may be involved in similar PAMP perception (Miya et al., 2007). This suggests that the perception of chitin or related PAMPs through LysM gene products may be widely distributed amongst plant species (Kaku et al., 2006). However, the predicted structure of CEBiP lacks intracellular domains which may indicate that additional components are needed to convey a signal to the cytoplasm through the plasma membrane (Miya et al., 2007). Co-operative function between LysM genes would not be unprecedented, as the *NFR1* and *NFR5* genes have been shown to act together in *Lotus japonicus* to recognize *Rhizobium* bacteria (Radutoiu et al., 2003).



Copyright © 2006 Nature Publishing Group
Nature Reviews | Molecular Cell Biology

Figure 1.2 A simplified model depiction of basal resistance activation following perception of PAMPs by PRRs. Model shows similarities between *R* gene and basal mechanisms such as the MAPK signalling and WRKY signal cascade. Figure also describes notable genes implicated to suppress basal defence through introduction of compounds through the Type III secretory system. Figure published by (Abramovitch et al., 2006a) with permission of Nature Publishing Group.

1.3.2 PAMP interactions with PRRs

The most widely-reported example of basal resistance involves perception of flg22 peptide, a highly-conserved bacterial peptide derived from flagellin, which activates basal disease reactions through a FLS2 receptor (Zipfel et al., 2004). FLS2 receptors instigate a signal cascade that is dependent on MAPKs and WRKY transcription factors to coordinate defence reactions including *PR* gene up-regulation (Schwessinger and Zipfel, 2008). The flg22 is detected directly by the extracellular LRR domain of the transmembrane receptor-like kinase FLS2. It has been demonstrated that flg22 is an effective elicitor when sprayed onto the surface of a leaf and presumably acts at an early stage of bacterial perception (Zipfel et al., 2004;

Zipfel and Felix, 2005). The FLS2 response has been associated with stomatal closures however *Pst* was shown to produce a polyketide toxin, coronatine, to initiate the re-opening of closed stomata in order to access plant tissues (Kwon, 2010). Bacterial exudates, which were refined to increase flg22 concentration, were sprayed on leaf tissue and also instigated basal defence reactions. flg22 has been shown to be the elicitor however it is likely to be one of many possible PAMPs capable of eliciting basal resistance through detection by LRR-RLKs (Zipfel et al., 2004). Binding is facilitated through formation of a complex with BAK1 that positively regulates the flg22 induced responses, and may have a signalling component through phosphorylation of another LRR-RLK, BRI1 (Chinchilla et al., 2007). BAK1 has recently been shown not to directly bind to flg22, it is therefore likely to be a signal transducer and general regulator of the interaction in which it may act in conjunction with FLS2 to confer specificity while adding to HR suppression (Schwessinger and Zipfel, 2008).

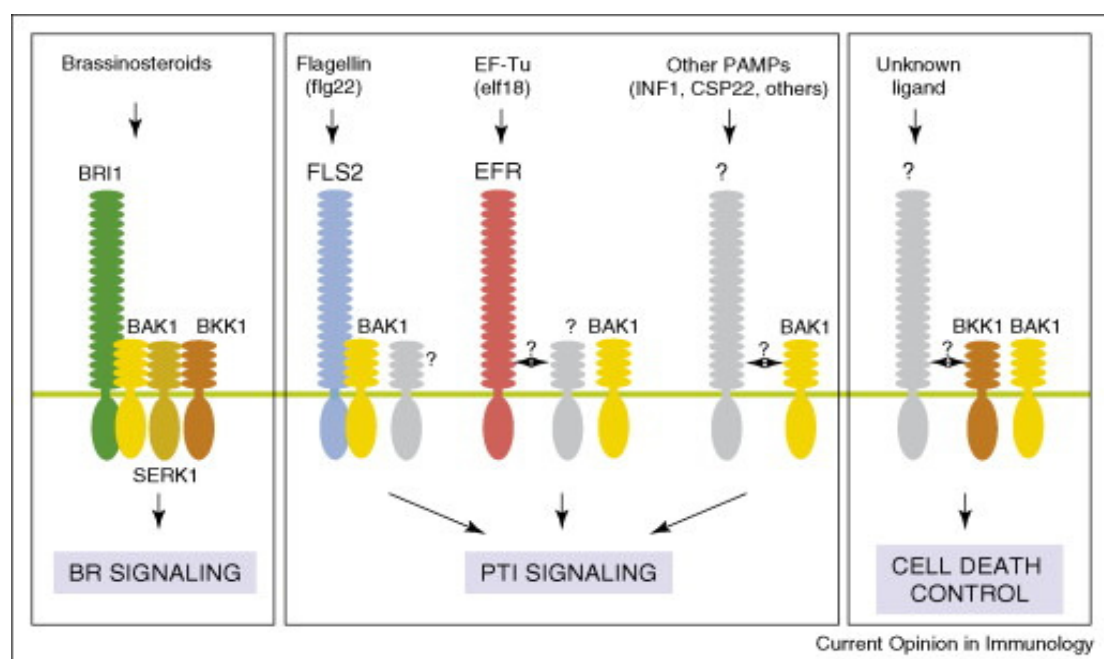


Figure 1.3 A simplified summary of the interactions between FLS2, BAK1 and other components in PAMP recognition. Figure shows the formation of a complex between FLS2 and BAK1 that result in signal transduction and basal defence. BAK1 and BKK1 have also been implicated as a regulatory facilitator in EFR PAMP perception and regulation of HR response which is typically suppressed during a basal response. Figure published by (Zipfel, 2008) with permission of Elsevier publishing.

1.3.3 Hormonal signalling in basal resistance

Some mutations affecting gene-for-gene incompatible interactions also affect basal mechanisms and *vice-versa*, which suggests strong mechanistic and regulatory conservation between the two defences. Transgenic plants, containing the bacterial salicylate hydroxylase gene *NahG*, effectively blocks SA tissue accumulation and compromises *RPS2*-mediated resistance. In addition, *EDS16* and *SID2*-disrupted genes prevent biosynthesis and create SA-deficient mutants (Fouts et al., 2003; Huang et al., 2005). These plants show enhanced disease susceptibility in both compatible and incompatible interactions, which indicates that basal disease resistance is mediated by SA accumulation (Fouts et al., 2003; Huang et al., 2005).

The JA-insensitive mutant *jar1* signifies JA as a component mediating basal responses as *jar1* shows enhanced susceptibility to the compatible *Pst* bacteria (Devoto and Turner, 2003; Nandi et al., 2003). Other mutants, such as *coi1* which disrupts the JA synthesis pathway, show similar increases in susceptibility to necrotrophic fungi *Alternaria brassicicola* (Thomma et al., 1998). In addition, the *fad3* mutant deficient in JA and the JA precursor linolenic acid acquire susceptibility to non-pathogenic species in wild-type plants such as *Pythium* (Staswick et al., 1998). This suggests that JA signalling is required for basal resistance and that basal resistance is likely to be involved in non-pathogenic innate immunity responses to non-host pathogens (Pieterse et al., 2001).

ET has been implicated in the regulation of basal responses as insensitive mutants exhibit enhanced susceptibility to compatible pathogens, such as *Pst*, and fungal pathogens including *Septoria* (Pieterse and van Loon, 1999). Mutants with reduced ET sensitivity show proportionally more severe symptoms when confronted with fungal pathogens, which suggests that ET is directly involved in basal response signalling. Basal resistance appears to utilise both SA-dependent and JA/ethylene-dependent pathways to produce an appropriate combination of responses to particular pathogenic threats (Pieterse et al., 2001; Nandi et al., 2003).

Hormonal signalling components typically associated with *R* gene-mediated responses, are widely shown to influence basal resistance activation to comparable extents (Glazebrook et al., 2006). This is consistent with the common hypothesis that much of the downstream mechanisms and regulatory pathways are conserved or utilised by *R-gene* mediated and basal defensive layers (Hammond-Kosack and Jones,

1997; Glazebrook et al., 2006; Jones and Dangl, 2006). Pathogen perception mechanisms and upstream signalling are likely to calibrate the extent to which downstream defences are mobilised and therefore define a resistance reaction. The components and function of these pathways require further investigation.

1.3.4 Basal resistance suppression

Two bacterial elicitors, *AvrRpm1* and *AvrRpt2*, have been shown to have an inhibitory effect on PAMP-induced signalling and therefore effectively suppress basal responses (Mackey et al., 2003; Kim et al., 2005a). The elicitors target the host *RIN4* gene product, which appears to be involved in basal PAMP signalling, and suppress its action (Kim et al., 2005b; Kover et al., 2005). Specific *R* genes including *RPS2* and *RPM1* appear to detect *RIN4* suppression when they are present and elicit an HR defence response, protecting the PAMP signalling pathway and illustrating the interplay between resistance pathways (Mackey et al., 2003; Espinosa and Alfano, 2004; Kim et al., 2005b). Other examples of Type III effectors that suppress basal defence are emerging in the literature including *AvrEI*, *AvrPto*, *AvrPtoB*, *HopAF1*, *HopAl2*, *HopC1*, *HopF2* etc. however the extent of their effect, especially on different genotypes, has yet to be fully defined (Espinosa and Alfano, 2004; Abramovitch et al., 2006a; He et al., 2007; Lewis et al., 2009). Plant PAMP-induced immunity is a target for pathogenic suppression and it is likely that many more components in such reactions are yet to be discovered and characterised (He et al., 2007).

1.3.5 Zigzag model

Specific examples of the complexity of the similarities and distinctions between pathways have been described above however, these interactions have been recently simplified into a “zigzag” model, which aims to illustrate and clarify our current understanding of the plant immune system (Jones and Dangl, 2006). This model shows the initial perception of pathogen-associated molecular patterns (PAMPs) by pattern recognition receptors (PRRs), resulting in a PAMP-triggered immunity (PTI) response which slows progression. The second stage describes effector-triggered susceptibility (ETS) in which the pathogen secretion of elicitors may interfere with

some aspect of PTI and increase virulence. Effector-triggered immunity (ETI) is described in the third stage where specific recognition occurs, either directly or through an intermediate, to bind to pathogen effectors and instigate an ETI (Jones and Dangl, 2006). This stage is commonly referred to as *R gene*-mediated resistance and is typically accompanied by localised, programmed cell death hypersensitive response (HR) if a resistance threshold is exceeded. The high selection pressure on pathogen population results in reduction of the targeted effectors within the population and so avoids ETI induction in stage 4 of the model. This stage predicts gain or modification to result in new effectors, which may be perceived by broad perception mechanisms, widely referred to as basal recognition (Jones and Dangl, 2006).

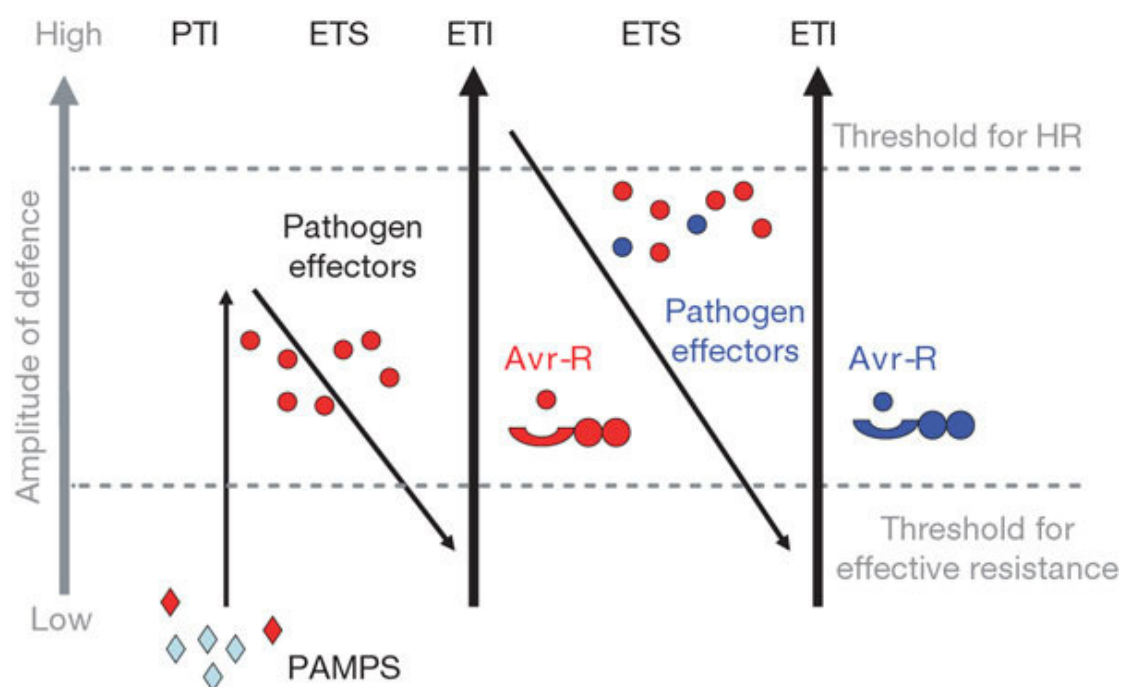


Figure 1.4 Model of plant immune system and its interaction with pathogens shows pathogen-associated molecular patterns (PAMPs) resulting in PAMP-triggered immunity (PTI). Effector-triggered susceptibility (ETS) is a result of suppression of components involved in PTI. However this can be overcome by specific recognition in effector-triggered immunity (ETI). Following these stages the model depicts pathogen loss of targeted effectors to increase virulence and plant evolution of receptors to counter this, as an example of defence component arms race. Figure produced by (Jones and Dangl, 2006) published with permission of Nature Publishing Group.

1.4 Natural variation

Gene function definition typically utilises phenotypic assessment of genetic variation within the population (Alonso-Blanco and Koornneef, 2000). Genomic studies assessing variation for a trait can provide a link between evolutionary analysis, natural genetic variation and molecular investigation (Mitchell-Olds and Schmitt, 2006). Variation in a complex multi-genetic trait, such as disease resistance, will alter throughout the population due to a wide range of influences including: gene sequence variation, expression variation, post translational modification, hormonal regulatory influence etc. However with use of genetic resources available in *Arabidopsis* (sequenced genome, genetic maps, genotyped accessions and QTL analysis), it is possible to take a system-wide approach to the trait to determine the predominantly influential factors. A typical example of such a study would be the assessment of natural variation in the *Arabidopsis* response to the *Avr* gene *hopPsyA* and its effect on the HR response. This investigation revealed two *Arabidopsis* loci significantly influencing control over the trait, with the highest LOD scoring loci mapped to a 22 cM region however, identification of the specific gene responsible is still to be established (Gassmann, 2005). Other studies detail the variation in light sensitivity or disease resistance compared with a variety of fitness scores in an attempt to correlate disease resistance to a distinct phenotype which, especially if combined with QTL data, can be easily Mendelised (Maloof et al., 2001; Kover and Schaal, 2002). Molecular isolation and characterisation of loci responsible for the naturally occurring variation has been achieved predominantly by using chromosome walking approaches and high-precision focused regional mapping (Alonso-Blanco and Koornneef, 2000). The addition of PCR based markers around the target loci aid fine mapping by reducing the search area however, such techniques require a strong QTL peak with a distinct phenotype so lines possessing the loci can be differentiated (Mitchell-Olds and Schmitt, 2006). An alternative approach, which uses transposon tagging within the QTL region to disrupt the functional effect, aims to isolate a specific responsible gene from candidates. This technique has been used successfully in maize to identify the *teosinte branched 1 (tb1)* gene, amongst others (Alonso-Blanco and Koornneef, 2000). When a specific gene has been identified it can be sequenced in various accessions to link molecular polymorphisms to known phenotypes as was successful for the *RPS2* gene (Caicedo et al., 1999). Natural variation screens are proving to be

an efficient alternative to mutant based approaches, which are often time intensive and have unpredictable success rates, especially in traits with weak alleles, pathway redundancy maintaining trait phenotypes or mutant lethality (Alonso-Blanco and Koornneef, 2000).

1.5 Economic importance of basal resistance

Pathogen infection is almost always characterised by a reduction in yield, although the method and severity of reduction is dependent on the disease strain. Yield reductions and crop losses are worth millions of pounds to the global agricultural industry and efforts to reduce disease using chemical treatments invariably increase production costs (Labanna and Banga, 1993). Consequently there is a high demand for disease-resistant crop varieties capable of producing high yields and offering broad protection against the majority of pathogens (Labanna and Banga, 1993). Most commercial variety breeding is still based around a conventional pedigree approach where morphological traits are ranked and assessed, with the highest scoring individuals being used to parent the next generation (Labanna and Banga, 1993). Once a commercially-appealing genotype has been developed, it is backcrossed several times until the plant is practically homozygous and the traits are stable across generations.

For much of the 19th century, traditional disease resistance breeding strategies had been focused on the inclusion of single *R* genes into commercial varieties from land races or other genetic sources. An *R* gene is race-specific and is very effective against certain strains but has little effect on others (Hammond-Kosack and Jones, 1997). Introduction of the *R* gene resistant species into a farm environment, grown at high densities and often in monoculture, will result in a large pathogenic selection pressure to overcome the plant resistance (Sprague et al., 2006).

There are numerous examples of new *R* gene-dependent disease-resistant varieties being introduced and their resistance being overcome within a few seasons (Sprague et al., 2006). Many breeders now focus on gaining resistance through polygenic partial resistance genes, which offer a less effective defence however are often broad-ranging and more durable in the agricultural environment. Partial or minor resistance genes are difficult to breed for and although their action is often not

fully understood, it is likely that their limited effect is a consequence of partial *Avr* gene recognition, resulting in a weak activation of defences (Hammond-Kosack and Jones, 1997). This is assumed because partially-resistant varieties are not effective against all pathogens, suggesting a broad-ranging but still pathogen-specific response, however the durability of effectiveness for such mechanisms will be dependent on the selection pressure they exert on the pathogen population (Hammond-Kosack and Jones, 1997).

Breeding for enhanced basal disease resistance may become a more widespread strategy as mechanisms are implied to be polygenic, which may improve resistance stability in a field environment if genetic markers can be reliably established. Basal resistance involves recognition of most viral, fungal, oomycete and bacterial races through detection of non-specific elicitors which are so fundamental to the infection process, they cannot be easily lost from an effective pathogen (Chisholm et al., 2006). Basal resistance may not be able to completely prevent pathogen proliferation however it typically slows pathogen development and progression, which would reduce widespread infection by slowing the pathogen lifecycle. The presence of high levels of basal resistance should elongate the pathogen lifecycle and so reduce overall pathogen population and pathogen colonisation rate in a field environment. Crop species with a high basal resistance are theoretically likely to be more resistant to epidemics as extending the pathogen lifecycle should reduce the period of optimal environmental infection conditions on which epidemics are reliant (Labanna and Banga, 1993). Breeding for effective basal resistance would mark a shift from traditional breeding strategies, which concentrate on making individual plants better at fighting disease, towards breeding for a population-wide control of disease (Hammond-Kosack and Jones, 1997). A strategy of enhancing basal resistance would concentrate on reducing total pathogen populations by slowing pathogen progression and extending pathogen lifecycles while accepting a reduced level of plant damage and loss to disease. Additionally, breeding for basal resistance may also prove effective in situations where *R* genes have not been discovered, have been overcome or do not exist. In an agricultural environment, a high level of basal resistance may prove an effective long-term and durable strategy for the reduction of disease and yield losses in crops.

PAMP recognition receptors, such as FLS2 and EFR have been shown to enhance pathogen resistance however transfer of such receptors to commercially

ready crop varieties has not yet been successful. However, successful transfer of the EFR PRR from *Arabidopsis* to *Nicotiana benthamiana* and *Solanum lycopersicum* has been shown to significantly increase resistance to a range of phytopathogenic bacteria (Lacombe et al., 2010). This example shows that heterologous PAMP component transfer between species is possible and may enhance the durable and sustainability of pathogen resistance in the field (Lacombe et al., 2010).

1.6 Experimental system

1.6.1 *Arabidopsis thaliana*

Arabidopsis thaliana is a member of the *Brassicaceae* family. This small (10-40cm) annual flowering species is found throughout temperate regions such as North America, East Africa, Europe and Asia and is commonly found on open, sandy soils which drain rapidly (Katagiri et al., 2002). The plant is used as a model organism despite having no commercial agronomic applications. However, synteny between plant species genomes allows research findings in *Arabidopsis* to be widely applied to economically important species (Gale and Devos, 1998).

Arabidopsis offers many advantages for research such as a small (114.5 Mb-125 Mb) diploid ($2n=10$) genome, which has been fully sequenced and has extensive genetic and physical maps associated with all five chromosomes (Glazebrook et al., 1997). There are relatively low levels of repetitive DNA, transposons, or other “junk DNA” between the 25,498 encoded functional genes which simplifies the genetics and increases the accuracy of predictions (Glazebrook et al., 1997). Most genotypes are highly productive self-pollinators (in excess of 10,000 seeds per plant) which allow recessive mutations to quickly become inbred and homozygous. Genetic mapping and mutation stacking can be easily achieved through cross-pollination (TAIR website). Its small size and rapid 5-6 week life cycle combined with its large mutant populations and efficient *Agrobacterium tumefaciens* mediated transformation protocols, have made *Arabidopsis* a very useful model organism for this project and other biological research (Katagiri et al., 2002).

1.6.2 *Pseudomonas syringae*

Pseudomonas syringae is a rod-shaped, Gram-negative bacterium which exists throughout temperate regions primarily following a saprophytic life cycle until it can successfully infect a variety of plant species, causing bacterial canker (Katagiri et al., 2002). The bacterium usually grows epiphytically using surface secretions as a carbon source in its early life cycle stages before initiating pathogenesis. Evolutionary pressure from plant defences have forced strains of the bacteria to become species-specific with complex methods of avoiding or neutralising antibacterial host products (Shan et al., 2007). If the plant is resistant to the *P.syringae*, then the infection is detected through race-specific elicitor binding to a plant receptor, usually bound to the plasma membrane (Glazebrook et al., 1997). This triggers defence reactions involving PR-protein production, reactive oxygen species (ROS), H₂O₂ induced lignification or cross-linking of the cell wall and the HR, which ultimately kills the invading pathogen (Dangl and Jones, 2001; Park et al., 2005).

In a compatible interaction with a host, the pathogenic *P.syringae*, which is spread by physical leaf contact or in rain splashes between infected plants, infects through wounded tissue or stomatal openings (Katagiri et al., 2002). The bacteria possess polar flagella that allow motility in the intercellular spaces where it proliferates and results in symptomatic chlorosis and water-soaked patches around the site of infection, followed by necrotic lesions and cellular collapse (Katagiri et al., 2002).

Several strains of *P. syringae* were found to infect *A. thaliana*, and such host-pathogen interactions have been used as a model pathosystem to investigate the molecular basis underlying plant disease resistance. Two virulent strains commonly used for research are *P. syringae* pathovar *tomato* DC3000 (*Pst*) and the *P. syringae* pv. *maculicola* ES4326 (*Psm*), which will readily infect *A. thaliana* when infiltrated into the intercellular space or applied to the leaf surface (Katagiri et al., 2002). This pathosystem has been extensively used to characterise plant defence reactions as symptoms develop quickly (1-3 days). In addition, the bacterial genome has been sequenced and also has an extensive array of genetic and physical maps associated with it. Through homologous recombination, the bacterium can be easily manipulated to include a wide array of reporter genes. This wide collection of advantageous properties makes *Pseudomonas* an ideal pathogen for use in this study.

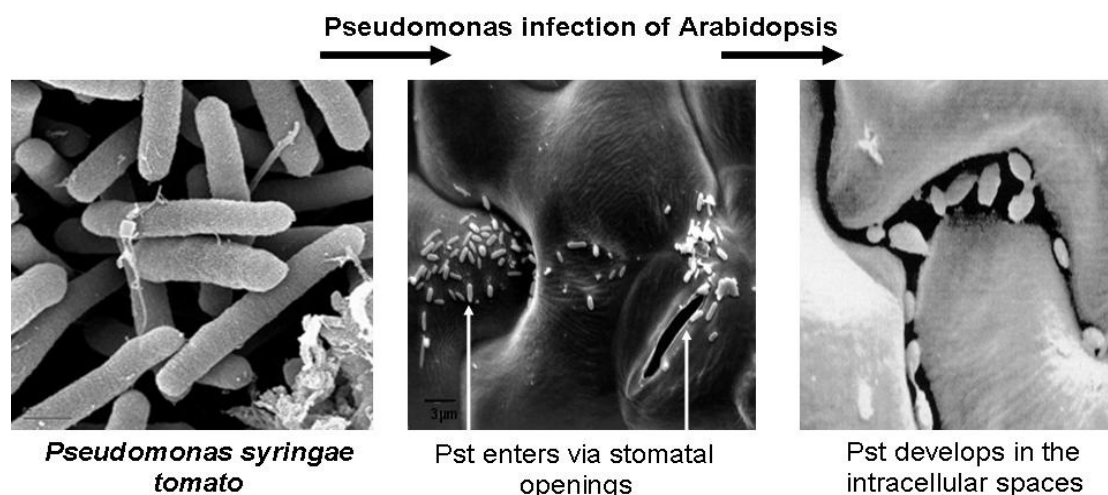


Figure 1.5 Electron microscopy images of bacteria infection process. Images produced at U.C. Berkeley and published with permission of the creator Gwyn Beattie.

1.6.3 Luciferase gene cassette

The interaction between plants and pathogens has been a powerful tool in determining the mechanisms behind R-gene mediated resistance (Mauricio et al., 2003; Jones and Dangl, 2006). Measurement of the bacterial growth *in planta* is routinely quantified by counting colonies of dilution plates from infected plant tissue. These procedures are time-consuming, arduous and difficult to scale for high-throughput assays. A newly developed assay utilises bioluminescent tagging of genes to bacteria, creating a quantifiable, definitive assay to provide pathogenesis data independent of symptom progression (Fan et al., 2008). The constitutively expressed *LuxCDABE* operon provides cellular components for generation of bioluminescence. Oxidation of luciferin emits a blueish-green light (wavelength ~560 nanometres) which can be quantitatively detected with a photon counter (Winson et al., 1998). The luciferase reaction is ATP-dependent and is maintained in the cytoplasm, which allows high-throughput *in vivo* assessment of cell viability and gene expression, through measurement of the light emission (Winson et al., 1998). Luminescence of the *LuxCDABE* operon has been surveyed in both *P.syringae* pv. *tomato* DC3000 and *P.syringae* pv. *maculicola* ES4326, concluding that the assay could detect small (1.3-fold) statistically significant differences in bacterial growth, which is comparable to plate assay precision (Fan et al., 2008).

1.6.4 Recombinant inbred lines

Recombinant inbred lines (RILs) are created through the repeated selfing of an F_1 hybrid to create near homozygosity at all loci in the RIL (Koornneef et al., 2004). Selfing increases homozygosity by 50% with each generation therefore, a segregating F_1 hybrid can be selfed to a practically homozygous genotype in eight generations (Young, 1996). RI lines have a theoretically non-segregating genome and so are effectively clones, which facilitate repeated phenotypic analysis on the same genotype. As the genome is effectively stable across generations, a marker map can be created using highly conserved domains and complementary markers to indicate the location of loci on the chromosomes (Glazebrook et al., 1997). Genotype information which assesses the parental origins of the recombinant genome sequence only has to be determined once for a RIL. RILs can be propagated eternally and are extensively used as tools for genetic mapping and assessing trait complexes (Young, 1996).

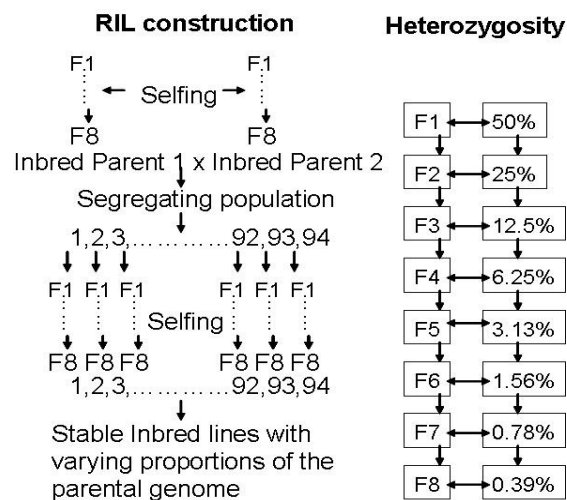


Figure 1.6 Flow diagram of RIL line construction and the reduction in heterozygosity between generations which results in practically homozygous stable lines.

1.7 Aims and objectives

Improving basal resistance in commercial lines will be facilitated by detailed understanding of the underlying mechanisms. Currently, this knowledge is only beginning to emerge in the literature and still requires extensive study before all components are known and characterised. It is the aim of this project to further this knowledge through assessment of the range of basal resistance variation naturally occurring within the *Arabidopsis* population. The project then aims to identify and characterise the genetic components responsible for variation in the basal resistance trait.

Preliminary results from other members of the group suggest that basal resistance may vary by up to 50-fold between ecotypes. The genetic and physiological basis for such resistance has not been extensively studied and has only implicated some possible downstream homology to SA-dependent genetic and signalling pathways (Veronese et al., 2003). Implications of upstream basal signalling and recognition of non-race-specific elicitor components are only beginning to emerge in the literature. Extensive identification and characterisation of the genetic, molecular and physiological components of basal resistance is an important prerequisite for any utilisation of this mechanism to improve crop plants.

This project will utilise and validate a recently developed, high-throughput luciferase-based assay, capable of quantifying pathogen concentration in excised leaf discs. The assay system has not been previously used to generate highly accurate and reproducible results. As a result, the interaction and procedures used initially need to be well characterised and then modified to refine the procedure to reduce replicate variation. Refinement experiments will examine the mean, and variation around the mean, from the alteration of a wide range of procedural components such as: bacterial concentrations, host and bacterial age at infiltration, a variety of environmental conditions, infiltration method, wound effect etc. Using this refined system, a wide screen of different genotypes and inbred populations will establish the range of variation in basal resistance.

Quantitative trait data, combined with the genetic maps and genotypic information available in specific *Arabidopsis* lines, allow for trait loci analysis which aims to indicate portions of the plant chromosome responsible for influencing a particular trait. Major QTLs identified in the first screen of a RIL population will be

compared against multiple other RIL lines to determine if major QTL locations are repeatedly high-lighted across independent populations. From this analysis, the genes contained in a chromosomal area that are statistically likely to influence the trait can be assessed for their properties, and a list of candidate genes responsible can be created. This is initially a similar approach that was used with great success by Caroline Dean's group and others in their investigation of the components involved in flowering time (Kowalski et al., 1994; Clarke et al., 1995; Laurie, 1997; Gazzani et al., 2003). An additional component would test a small range of different virulent pathogens on RIL populations to indicate the relative strength of basal defence in comparison with other ecotypes. Comparison between quantitative trait loci (QTL) profiles generated from different pathogens would help to indicate the likelihood of whether the responses are broad or pathogen-specific.

QTL analysis should define the number and location of major QTLs across the *Arabidopsis* population which substantially influence basal resistance. From the QTL locations, a list of candidate genes should be produced that have potential for involvement or previous implication in disease recognition and prevention pathways. Molecular characterisation and transcriptional analysis would prove useful in characterising the potential extent of the candidate genes' action and possible mechanistic role in basal resistance. Direct proof of a candidate's involvement then needs to be established to complete the project aims of identifying the genetic components involved in basal resistance.

Chapter 2: Materials and methods

2.1 *Arabidopsis* materials and methods

2.1.1 *Arabidopsis thaliana* genotypes

Seeds for a set of 92 inbred parental genotypes of recombinant inbred lines (RILs) were obtained from a collection held by Claire Lister (John Innes Centre, Norwich, UK). Using trial data collected from these genotypes, a set of 5 RILs were selected from crosses between Sorbo x Gy-0, Nok-3 x Ga-0, Cvi-0 x Ag-0, Col-0 x Ler and Wt-5 x Ct-1. For each of these crosses, 94 RILs seeds were obtained from the collection generated by Ian Bancroft's group (John Innes Centre, Norwich, UK) and 29 SALK T-DNA insertion lines were ordered from Nottingham *Arabidopsis* Stock Centre (NASC).

2.1.2 Plant growth

Prior to growth, seeds were stored in dry microfuge tubes at room temperature in the Laboratory. Seeds were stratified by immersion in 0.1% agarose solution in microfuge tubes for 3 days at 4 °C. Seeds were distributed with a glass pipette into pots containing “*Arabidopsis* mix” (a 12:1 mixture of Scott's M2 plus Intercept to grit). Plant pots intended for disease trialling were placed in a short day growth room for 4-5 weeks at 9 hrs day: 15 hrs night, at ~23 °C, ~70% relative humidity and light level at between 100 to 125 $\mu\text{mol m}^{-2} \text{s}^{-1}$ (shelf level). Trays were covered with a clear polystyrene lid after sowing for 4 days to promote germination and watered regularly from underneath to avoid seed disturbance.

Plant intended for seed bulking or crossing were grown in glasshouse / long day growth room conditions for 6-8 weeks at 22 °C, 65% relative humidity, photoperiod 15 hours with natural glasshouse light levels supplemented with Osram 4Y 400w Vialox Nav-T super bulbs to maintain day length.

Seeds for media plate growth were sterilized by washing with 70% ethanol for 5 minutes followed by 20 ml 5% v/v bleach (Vortex) + 0.1% SDS for 20 minutes. The bleach was aspirated off in a flow hood and the seeds were washed with two changes of sterile distilled water and placed on MS + 0.6% agarose plates.

2.1.3 Media composition

KB (Kings B Medium)

Formula per 1 litre de-ionised water

Proteose Peptone 20g/l

Glycerol 1.6g/l

pH to 7.2 with NaOH

For solid 15g Lab M agar

LB (Luria-Bertani) – G broth and LB – G agar

Formula per 1 litre of de-ionised water

Tryptone 10.0g

Yeast Extract 5.0g

NaCl 10.0g

Adjust to pH 7.0 with 1M NaOH

For solid medium add per litre

Lab M No.1 agar 10.0g

2.2 Pathogen materials and methods

2.2.1 *Pseudomonas syringae* storage and growth

Luciferase reporter gene-tagged *Pseudomonas syringae* pv. *tomato* DC3000 and *Pseudomonas syringae* pv. *maculicola* ES4326 strains were provided by J. Fan (JIC, UK). *Pseudomonas syringae* pv. *tomato* DC3000 *AvrPtoB* deletion mutant was provided by C. Zipfel (Sainsbury Laboratory, UK). Long term pathogen stocks were held in 50% glycerol at -80 °C before transfer to agar stock plates on KB medium supplemented with 50 µg/ml each of rifampicin and kanamycin, stored for up to 20 days at 4 °C. Bacteria grown for disease trials were spread onto a Petri-dish of KB media with appropriate antibiotics and incubated at 28 °C overnight (12-16 hours) to obtain a confluent lawn. To ensure reproducible results for each of the parental ecotypes and RIL lines, four replicate plants were grown and three infected leaf discs were taken from each of the host plants (minimum 12 data points per RIL).

2.2.2 *Pseudomonas* infection

The bacterial inoculum was prepared by taking an aliquot from a fresh lawn of bacteria (12-16 hrs) and suspending it in 10mM MgCl₂. The bacterial density was measured using an Eppendorf Bio-photometer and adjusted to OD₆₀₀ 0.002

(equivalent to $\sim 10^6$ cfu/ml) with 10mM MgCl₂. Pressure infiltrations were made using 1 ml needleless syringe on the abaxial side of the leaves so as much of the leaf as possible was filled with inoculum (so that the leaf appeared water-soaked). Dipping, vacuum and spray infiltrations were facilitated by the addition of surfactant (0.02% Silwet) to the suspension before direct application or aerosol spraying of the suspension onto all leaf surfaces. Plants inoculated using methods other than infiltration were covered with plastic lids to raise humidity and aid infection. Following infection, plants were returned to the short day growth room, for approximately 48 hours. Bioluminescence of 8mm leaf discs was measured using a Berthold FB12 luminometer with single photon counter (Fan et al., 2008). Data is recorded as photon counts per second (cps) per disc. To assess bacterial content in non-Lux tagged strains or in comparative trials, dilution plating and colony counting was used by macerating infected tissue in KB medium. Serial dilutions of between 10^0 to 10^{-6} were spread onto KB plates and incubated for 2 days at 28°C for colony counts. In certain experiments, infected leaves were also given a visual score of disease severity (1=low, 5=high) before quantitative assessment.

2.2.3 Infiltration controls and standardisation calculations

To improve reproducibility from initial trials, luminescence values from leaf discs were compared against internal controls. Three data points per plant were averaged when collected and replicate genotype data from other trays was also included into a single average value. A percentage comparison against a set of 4 Col-0 controls in every tray as well as 4 controls of a parental line, was used to reduce tray-to-tray variation.

2.2.4 Refinement of technique experiments

The Lux-tagged *Pseudomonas* high throughput assay is a newly-developed system and not yet fully optimised and so preliminary data was highly variable. A series of experiments was conducted to characterise the infection process and optimise the assay so that reproducible data could be collected, which is an important prerequisite for accurate QTL analysis. Moderate alterations were made to various aspects of the

system and their effect on the trait and variation from the mean were observed and used to refine the assay. Alterations included: time course of infection, density of bacterial inoculum and bacterial population age at time of infiltration e.g. 6, 12, 18 hrs. In addition, plant age and time of day of the inoculation were assessed. Furthermore, different methods of inoculation were assessed for reproducible infection and alterations to various environmental factors were tested. Modifications to the original procedure that reduced variation and improved reproducibility were introduced as standard for subsequent trials. Statistical analysis to verify the effect of a treatment typically utilised the 'Genstat' 10th addition statistical software from VSN International. For treatment effect analysis the majority of this project utilised one-way analysis of variance (ANOVA) e.g. section 3.4.3 however, regression analysis was used to test for a proportional relationship between variable e.g. section 3.2.1. The F-test, or variance ratio test, was used to compare standard deviations between groups and test the null hypothesis that two populations have the same variance, producing a P value showing significance of the variance ratio e.g. section 3.4.3.

2.2.5 Parental RIL line screening and selection of appropriate RIL lines

Following refinement of the assay technique, a range of 92 genotypes, which had been used as parents of the associated RIL line populations, were selected and screened for their natural variation in the basal resistance trait. A continuous range of variation in the basal resistance was observed with an approximate 50-fold difference in bacterial growth between the most and least susceptible accessions. Examination of bacterial growth in the parents of the RILs allowed the selection of 5 populations which showed different levels of variation so that as many significant QTLs as possible could be found responsible for altering the trait. Gy-0 x Sorbo, Ga-0 x Nok-3, Wt-5 x Ct-1, Ag-0 x Cvi-0, Ler x Col-0 lines were selected, each comprising at least 94 RILs, were then surveyed to provide the quantitative trait data for subsequent QTL analysis. Subsequent comparison of QTL locations could then be used to determine if the variation at high susceptibility was influenced by the same QTLs as variation at relatively low susceptibility. From these choices of RIL populations, an understanding of how many QTLs are involved in basal resistance may be elucidated.

2.2.6 Visualising photon emission from Lux tagged bacteria using NightOWL luminescence camera

P.syringae transformed with the Lux gene cassette emits photons at ~560nm which act as a reporter of viable bacteria within plant tissue. Berthold Technologies have produced the “NightOwl” system, which can detect this photon wavelength and density using a slow-scan CCD camera in a light-proof chamber. Newly-infected plants were placed in the cabinet and over a 5-day infection period, a series of single frame images were taken. Col-0 and Ler genotypes were used and both *Pst* and *Psm* were used for comparative infiltrations. Images used a 15-minute exposure time, taking 24 pictures per day to create 120 sequential images in total. Photoshop CS3 software was used to standardise image size and ensure correct alignment and orientation. No alteration was made to the contrast, colour, image quality, etc. ‘ImageJ’ software was used to stack image sets of sequential JPEG files and compile them into movie MPEG files similar in composition to the time-lapse animation technique. The movie files run for 30 seconds at 4 frames per second.

2.3 QTL analysis

QTL analysis can be used to explore the genetic components underlying the natural variation of basal disease resistance. The analysis enables tracking of the chromosomal regions that influence a trait. A variety of statistical tests are run to ensure consistency in the analysis and reduce the possibility of statistical artefacts which are uncommon, but possible (Van Ooijen and Maliepaard, 1996; Young, 1996).

Computational statistical analysis assesses the effect of marked genomic regions on the phenotypic trait and generates a probability of association between DNA marker and trait (Liu, 1998). Repeated nonparametric analysis is used to assess the data as it makes no assumptions about the data being of normal distribution. QTL mapping software rank all the marker and genotype data with reference to the trait information based on the variance from the null-hypothesis (Falconer and Mackay, 1996; Liu, 1998). A marker linkage group with a segregating QTL will show a statistical gradient towards the locus with the greatest influence on the trait (Van Ooijen and Maliepaard, 1996). The software interpretation of this gradient is a set of line graphs covering the length of each chromosome examined. The line measures the

probability of a locus on influencing the trait and often calculated as logarithm of odds (LOD) score. An ideal QTL peak has a high LOD score over a tightly-defined chromosomal area.

2.3.1 QTL analysis information

Marker and genotype information for the RI lines associated with Col-0 x Ler cross were obtained from C. Lister (JIC, UK) (Dean et al., 1991). These Restriction Fragment Length Polymorphism (RFLP) maps were combined with publicly-available data e.g. NASC (<http://arabidopsis.info/>) TAIR (<http://www.arabidopsis.org/>), consisting of PCR markers such as: Simple Sequence Repeats (SSR), Random Amplified Polymorphic DNA (RAPD), Amplified Fragment Length Polymorphism (AFLP) and Single Nucleotide Polymorphism (SNP). As a result, a high-density marker map was created with an excess of 1300 DNA markers, at an average interval of 0.6 cM (centimorgan), which has been genotyped in 100 RILs to determine parental origin at the marker location. Similar information, but at a lower marker density, was donated by I. Bancroft's group for the remaining 4 sets of 94 RILs (http://www.jic.ac.uk/staff/ian-bancroft/research_page5.htm). Quantitative basal trait data was compiled from at least 12 data points per RIL into a single value describing relative disease susceptibility of the line.

2.3.2 Linkage mapping methods

QTL analysis was performed using 'MAPQTL5' software (Van Ooijen and Maliepaard, 1996) where Kruskal-Wallis, interval mapping, restricted multiple QTL mapping (MQM) and non-restricted MQM analysis were calculated. Automatic cofactor selection (ACS) was used to highlight cofactors to the QTL and permutation tests (PT) were performed to calculate the 95% max log-likelihood (LOD) score for both chromosome-specific and genome-wide statistical significance. The majority of QTL results shown are derived from MQM analysis as it is the most robust method. However other tests offer a useful comparison and increase confidence in the validation of the final result.

2.3.3 QTL analysis software and files

'MapQTL 5' QTL analysis software from the 'Kyazma' company was used which requires 3 file types containing the relevant information presented in a specific format. A master .mqd file is initially generated which will include the analysis but this must be accompanied by a .loc file which contains the genotype data. Markers and their chromosomal locations are described in a .map file and the quantitative trait data is described in a .qua file. The data used for the .map and .loc files is available on the Nottingham *Arabidopsis* Resource Centre (NASC) website (http://arabidopsis.info/new_ri_map.html) or from the Bancroft group (JIC, UK) website (<http://www.jic.ac.uk/staff/ian-bancroft>).

2.3.4 Kruskal-Wallis QTL analysis

To give a basic analysis comparison, the analysis methods were firstly run without co-factor selection. Once the data is loaded into the MapQTL programme, a Kruskal-Wallis test (Lehmann, 1975) is used to rank all individual RIL genotypes according to their trait value and parental origin genotype information. The Kruskal-Wallis analysis gives a basic comparison of genotypes and trait scores between the RIL lines without any specific loci being screened for co-factor selection. QTLs of large effect will be highlighted by big differences in average rank of the marker genotype classes. The QTL test statistics rank the genotype classes and gives a chi-square style distribution. A distinct change in the test statistic compared to background values indicates the locus with the closest linkage to a QTL, confirmed by the significance level. However the Kruskal-Wallis test is limited as individuals with either marker location, genotype data or a quantitative trait score missing will be removed from the analysis, which reduces the power of the test.

2.3.5 Interval mapping QTL analysis

Interval Mapping (IM) was used to test whether QTLs at different positions have a genetic effect. This creates a likelihood ratio, transformed to Log^{10} to give the log of odds (LOD) score, which indicates the potential sites of QTLs. A high LOD score suggests the presence of a QTL but to find the threshold above which the LOD value becomes significant, a permutation test is required (Churchill and Doerge, 1994).

2.3.6 Permutation tests in QTL analysis

This test looks at the frequency distribution of the maximum LOD score for each linkage group and over the whole genome. After 1000 permutations, an estimate of the distribution of the LOD deviance under the null-hypothesis that there are no QTLs is given. By looking at the genome-wide scores, it is possible to find which LOD score represents the 95% confidence interval (typically above $\text{LOD} = 2$) which is then used in subsequent analysis to locate QTLs.

2.3.7 Restricted MQM mapping and MQM mapping QTL analysis

Multiple QTL Modelling (MQM) is then used where, unlike IM, the QTLs are identified due to having a LOD score on or above the threshold. This is determined by the permutation test where co-factors can be included in to the model, to lower any residual variance (Arends et al., 2010). MQM mapping firstly augments and models multiple genotypes to estimate their comparative probability of trait influence then important markers are selected by multiple regression analysis and backwards elimination of lower probability markers. The QTL is then aligned to the chromosome with reference to genotype information and co-factors of influence. MQM mapping has advantages over IM as it is a more stringent test, compensating for the residual variance from reference to the full analysis model to produce more accurate peaks and is less likely to produce statistical ghosts (Arends et al., 2010). After the first MQM, the original IM QTLs are more defined and any new ones become visible or significant. MQM mapping is an important and stringent test so these results are observed most carefully.

2.3.8 Automatic Co-factor Selection in QTL analysis

After an initial run of MQM analyses to get a benchmark QTL profile, Automatic Co-factor Selection (ACS) is used to assess large numbers of potential co-factors to represent each linkage group. Assessment of multiple-QTL models is more powerful than single-QTL model predictions however computational complexity increases dramatically. Marker regions with a smaller proportion (typically non-significant) of influence on trait variation are defined as co-factors to the larger (typically significant) chromosomal areas of influence on the trait. Co-factors function as a genetic background control and statistically reduce fluctuation of the residual variance from nearby QTLs (Van Ooijen, 1999). The resulting accuracy of QTL analysis is enhanced and the resolution of significant QTL peaks is greater while low scale genetic flux influence over the trait is minimised (Van Ooijen, 1999). A process of backwards elimination is then used to define marker regions that contribute the highest amounts of variation explained by the model, thereby narrowing the regions containing QTLs and resolving QTL peaks to show significance. Following the selection of appropriate co-factors Kruskal-Wallis, interval mapping, restricted MQM, MQM mapping and the permutation tests are re-run to a more accurate standard. The final QTL graph outputs are produced with similar profiles to the previously generated analysis but with a higher resolution.

2.4 Candidate gene identification and refinement

Markers located on the peripheral edges of QTL peaks were selected and the genomic DNA sequence region was isolated using the Sequence Viewer tool on the TAIR website (<http://www.arabidopsis.org/servlets/sv>). Genes contained within those regions were listed and compared against a list of known disease-related or PAMP-associated genes, derived from 'MapMan 2.0' microarray analysis software and flg22 treatment microarray data (Zipfel et al., 2004). The cross-referencing of these lists of genes and further comparison between the independent QTL analyses of separate *Arabidopsis* populations allows for an approximate 95% reduction in the total number of candidate genes. Detailed examination and characterisation of the genes and gene groups within this remaining candidate list was used to determine if any candidates had the potential to be involved in the basal disease resistance pathway.

2.4.1 Candidate and disease related gene lists

Chromosomal markers either side of significant QTL peaks were selected to determine physical genetic locations for QTL peaks, calculated in centimorgans. Using the web-based TAIR Sequence viewer (<http://www.arabidopsis.org/servlets/sv>), the genetic markers can be entered to isolate the chromosomal region within markers. The software can then generate a list of all genes within the region which then forms the basis for the candidate gene lists for each significant QTL peak. All the genes within the regions are listed, without assigning probabilities, according to location and peak height.

Microarray software 'MapMan2' from 'Soxoft' contains a hierarchical organisation and categorisation of genes with reported function. From this software, a list of 421 genes was compiled, reportedly involved in biotic stress based on the gene ontology from currently published annotation information available. The biotic stress list was combined with a list of 269 genes reportedly involved in PAMP recognition based on a microarray transcriptional analysis when *Arabidopsis* is exposed to a flagellin flg22 protein exudate (Zipfel et al., 2004). The combination of these 2 gene lists creates a record of currently identified disease /PAMP related genes.

2.4.2 Refinement of candidate lists and categorisation of genes

QTL-derived candidate gene lists can be cross-referenced against the disease related gene lists using a multiple interrogation query in 'MS Access' software. Refined candidate gene lists, containing only genes reportedly involved in disease resistance and located within each of the total QTL regions detected, can then be generated by appropriately removing candidate genes. Genes contained within the refined candidate gene lists can be categorised based on their annotation information into gene families of related properties for further analysis.

2.4.3 Candidate gene distribution within QTL regions

Cross comparison of the categories of candidate genes contained within QTL regions, combined with statistical analysis, can determine if the frequency of a particular gene type/ family is significant compared to the background distribution. Chi squared matrices compile information including the total gene number, number of genes in the

category of interest, total number of genes found within the QTL regions and number of genes of interest within the QTL regions. From this information and using 'R-2.6' statistical programming language to compute the Chi squared matrices, it is possible to compare the effective hit ratios against the background distribution and prove that the gene of interest is occurring at a significantly higher frequency within the QTL region.

2.5 LysM gene characterisation

2.5.1 Annotation information

Sequence and annotation information of the *Arabidopsis* LysM genes were obtained from the TAIR website (<http://www.arabidopsis.org/>) which lists currently known or implied function and properties from the latest published research literature and community data submissions.

2.5.2 Phylogenetic analysis

The nucleotide and amino acid sequences of the LysM members were used to generate phylogenetic trees to visualise the coding similarity and evolutionary relationship. Diagrams are generated using sequential combination of outputs from: 'Muscle', 'Genedoc', 'ClustalW', 'Phylip 3.67', 'Protdist', 'Neighbour', 'Consense' and 'Drawgram'. This software, available on the JIC intranet servers, utilises bootstrap analysis and distance matrix methods to produce evolutionarily robust phylogenetic trees (Baldauf, 2003; Harrison and Langdale, 2006). The 'Align X' software tools from the Invitrogen 'Vector NTI' package was also used to produce tree diagrams with combined sequence identity, absolute complexity data and gene matrices showing the gene similarity percentage (Baldauf, 2003).

2.5.3 Expression profiles of LysM genes

Amalgamated Affymetrix microarray data from a large variety of biological contexts has been combined in an internet based software tool; 'Genevestigator V3'

<https://www.genevestigator.com>). This powerful meta-analysis algorithm provides tools to explore specific gene expression from a developmental, stimulus response, mutation or anatomical perspective (Hruz et al., 2008). Modes of expression can be visualised through creation of heat maps illustrating gene responses to stimuli and profiles indicating spatial and temporal gene expression. With reference to individual expression databases, a model of the gene role within the context of metabolic and regulatory networks can be hypothesised (Hruz et al., 2008). Expression profiling involves use of meta-profile analysis, biomarker search, pathway projector and clustering analysis for compiled candidate gene lists within the software and facilitates output of expression graphs and ranges. While ‘Genevestigator’ is a useful tool to formulate hypotheses and characterise gene expression it must be considered that the database is compiled from amalgamated data therefore without specific reference to experimental data, the trends observed may be open to other interpretation. ‘Genevestigator’ however, does allow access to the source experimental components of the amalgamated data so it is possible to access specific experiments which may be more relevant within the basal resistance biological context.

2.5.4 Sequencing

DNA sequencing was done in collaboration with V.Lipka’s group (SL, UK) to assess the extent and effect of polymorphisms in candidate genes for selected genotypes. Following PCR amplification of a candidate gene sequence, the DNA from different genotypes was submitted for automated preparation and analysis at the JIC Genome Laboratory using ‘AbiPrism 3730XL’ capillary sequencers and ‘ABI’ software to provide comparative nucleotide sequence data. An existing protein sequence data set containing sequences for the majority of candidate genes in an array of genotypes was obtained from D.Weigel’s group (Max Planck, GER). Nucleotide sequence for the fully sequenced Col-0 was obtained from the TAIR website and Ler genotype sequence data was provided by Monsanto. Sequence alignments and comparisons were analyzed using Align X and other tools from the Invitrogen ‘Vector NTI’ package.

2.6 T-DNA knockout line identification

2.6.1 T-DNA insertion lines

T-DNA insertion lines were ordered from the NASC (<http://arabidopsis.info/home.html>) and they were predominantly chosen to have originated from a Col-0 background. Lines selected preferentially contained an insertion sequence within the exons of the LysM genes. However where such lines were not available, lines with insertions within 300 bp upstream of the start codon, downstream of the stop codon or within the intron region were selected. Small numbers of plants of each T-DNA insert lines (minimum 16) were grown and analysed for disease resistance, as in the QTL trials, with 8 controls per p24 tray for comparison, as with previous experiments. Where possible, homozygous lines were selected however, for many of the lines, the insertion T-DNA is still segregating so the homozygosity of the T-DNA must be checked by PCR. Following bacterial growth trait scoring, a sample of leaf tissue was taken from each plant and its DNA extracted and purified. Specifically designed primers for each of the knockout lines were used to set up two paired reactions using three primers (LP, RP LB insertion). From the electrophoresis gels run, differences in amplified DNA fragment length pattern between paired reactions indicated insertion genotype as described in Figure 5.7. Repeat PCR runs were conducted wherever possible and the results cross-compared to increase reliability until a comprehensive list of plant insertion genotypes could be produced, accompanied with trait data for each plant of the knockout lines.

Bacterial growth tests were conducted with the infiltration of *Pst*, following the procedures described in Section 3.3.1 and 3.3.2. A minimum of 16 plants were tested for each insertion line however, following PCR testing to verify insertion type, the sample number is variable as most lines are still segregating for the insertion and so typically tested as a mixture of heterozygous, homozygous or wild type.

2.6.2 DNA extraction

Arabidopsis young leaves frozen in liquid N₂ were ground and DNA was extracted with an extraction buffer (250mM NaCl, 200mM Tris-Cl, 25 mM EDTA, 0.5% (w/v) SDS) and centrifuged to separate cell components. DNA samples were washed with isopropanol (100%) and ethanol (70%) sequentially with centrifuge cycles at 13,000 g

for 5 mins intervals. The DNA was then resuspended in 100 µl of 1x TE for the stock and a 100 fold dilution was used for PCR reactions. Stocks were kept in a -20°C freezer.

DNA quantity and quality was assessed using an Eppendorf Bio-photometer UV setting to measure absorbance at 260 and 280 nm. Comparative absorbance ratios between A_{260}/A_{280} were calculated with a ratio of 1.6 to 1.8 being considered an acceptable level of DNA purity. $OD_{260} = 1$ is equivalent to 50 µg/ml of double stranded DNA.

2.6.3 Oligonucleotide primers

Primer sequences were designed according to standard specifications using Invitrogen ‘Vector NTI Version 10’ or with the web-based SALK T-DNA primer designer (<http://signal.salk.edu/tdnaprimers.2.html>). All primers were checked to ensure accurate genomic placement. All synthetic oligonucleotides were produced by Sigma-Genosys/Aldrich (<http://www.sigmaaldrich.com/life-science/custom-oligos/custom-dna.html>). Dried oligonucleotides were resuspended in 1 x TE (10mM Tris, pH 8.0, 1 mM EDTA) to produce stock concentration of 100 µM. Working solutions for PCR reactions were diluted 10 fold in sterile water to give a concentration of 10 µM and all primers were stored at -20 °C.

2.6.4 Polymerase Chain Reaction

Polymerase chain reaction (PCR) cycling conditions were optimised for each primer pair using technical data calculated by Sigma-Genosys. PCR reactions were performed in a PTC-200 Peltier thermal cycler from MJ Research with an extension time of 1 min / 1 kb length of PCR product. Components of a 10 µl reaction are listed as following:

• DNA template	1 µL
• 10x Buffer 500mM KCl, 100mM Tris-HCl (pH 9.0), 1% Triton X-100, 15mM MgCl ₂	1 µL
• dNTPs (25mM each dNTP base– Invitrogen)	1 µL
• Forward Primer 10µM	1 µL
• Reverse Primer 10µM	1 µL
• Taq (Amplitaq gold 5u/µL – Applied Biosystems)	0.25 µL
• dH ₂ O	4.75 µL

Some reaction components were adjusted for refinement purposes.

Thermal cycling programmes were set up as follows:

1 x Cycle	Initial denaturation	5 min	94°C
30 x Cycle	Denaturation	1 min	94°C
30 x Cycle	Annealing	1 min	58 °C
30 x Cycle	Extension	1 min/ kb	e.g. 72°C (Primer T _m)
1 x Cycle	Final extension	10 min	72°C
1 x Cycle	Hold	Overnight	8°C

Some reaction components were adjusted for refinement purposes.

2.6.5 Agarose gel electrophoresis

PCR products were separated by electrophoresis in a submerged 1% (w/v) agarose gel prepared with 1 x TBE buffer (90 mM Tris-borate, 2 mM EDTA) (pH approx 8.0), 0.1 µg/ml ethidium bromide. Loading buffer (10 mM Tris-HCl, 1mM EDTA, pH 8.0, 50% (v/v) glycerol, 0.05g/ml Bromophenol blue) was mixed with the PCR products and the marker ladder was loaded into an adjacent lane. The gels were run at 10-20 V/cm for 45 mins and the DNA was visualised with UV light (302nm).

Chapter 3: Refinement and characterisation of the luminescence-based *Pseudomonas* growth assay using *Arabidopsis* ecotypes

3.1 Introduction

For centuries, plant breeders have exploited the range of natural variation in phenotypic traits to select properties which are advantageous to cultivation in a simplistic but effective strategy of breed best with best. Modern methods of crop improvement aim to understand the genetic basis of the trait and causes for its variation across the population so that targeted, and sometimes indirect, cross-breeding results in agronomically superior varieties (Asins, 2002). Assessment of natural genetic variation between *Arabidopsis* accessions and *Pseudomonas* isolates has proven a successful approach in the identification of *RPS2* and *RPM1*, members of the NBS-LRR class of *R* genes (Aarts et al., 1998). Substantial natural genetic variation was also found when the *RPS2* gene was sequenced in 27 *Arabidopsis* accessions, which results in protein product variants that correlate with alterations in resistance efficacy to *Pst* (Mauricio et al., 2003). A major component affecting natural variation in resistance to *Peronospora parasitica* across the *Arabidopsis* population was traced to the *RPP5* gene. The encoded protein is a member of the TIR-NBS-LRR class and diversity in its structure between *Arabidopsis* ecotypes such as LRR length has been linked to alterations in host specificity for several bacteria (Parker et al., 1997). Natural variation of *R* gene structure has been demonstrated to significantly affect resistance and has led to detailed mechanistic characterisation and understanding of *R* gene function, therefore, similar characterisation of natural variation may be successful in determining the genetic basis for basal resistance.

Basal resistance is often exemplified by the ability of a plant to limit the pathogenesis of virulent pathogens in the absence of *R-avr* gene interaction. Therefore, *Pseudomonas syringae* pv. *tomato* DC3000, that has been shown to be universally virulent to *Arabidopsis* accessions, was chosen for this study (Fan et al., 2008). Preliminary use of the high-throughput quantitative Lux-tagged *Pseudomonas* assay, to measure bacterial content *in planta*, showed a wide range (~50 fold) of variation in basal resistance across a selection of *Arabidopsis* ecotypes (Jun Fan, pers. comm.). The Lux-tagged assay measures bacterial number as a direct correlate to photon emission, which produces an unambiguous gauge of basal resistance with which to test ecotypes of the population. The assay was newly-developed and largely

uncharacterised, and therefore a set of experiments were designed to optimise the procedure, reducing replicate variation to facilitate QTL analysis. Alterations were made to various aspects of the pathosystem, as described in Table 3.1, and their effect on bacterial growth and the experimental reproducibility among replicates was observed and used to refine the assay. Lux-tagged *Pst* photon emission was used to visualise the infection progression and qualitatively characterise bacterial proliferation rate and extent within the tissue.

3.2 Results

3.2.1 Relationship between Lux emission and bacterial number

For the newly developed Lux tagged bacteria assay to be useful as a reporter of viable bacteria in plant tissue, it is necessary to determine the relationship between light emission and bacterial number *in planta*. Different lines from the Sorbo x Gy-0 RIL population were initially infiltrated with the same concentration of Lux-tagged *Pst* (OD₆₀₀ 0.02). Data points were derived from the average of 3 counts per second (CPS)/disc values and 3 dilution plate count values per accession, which showed a range of variation in bacterial content and luciferase readings. Regression analysis compared the luciferase luminescence measured in CPS with bacterial number measured in colony forming units (CFU) determined by the plate assay. The results show a significant positive correlation ($P = <0.05$) between luminescence and bacterial number, which confirms that the Lux assay generates an accurate representation of pathogen content in the plant tissue. The R^2 value of 0.897 shows a percentage explanation of the data based on the line of best fit, and so almost 90% of the variation observed is explained by a linear equation.

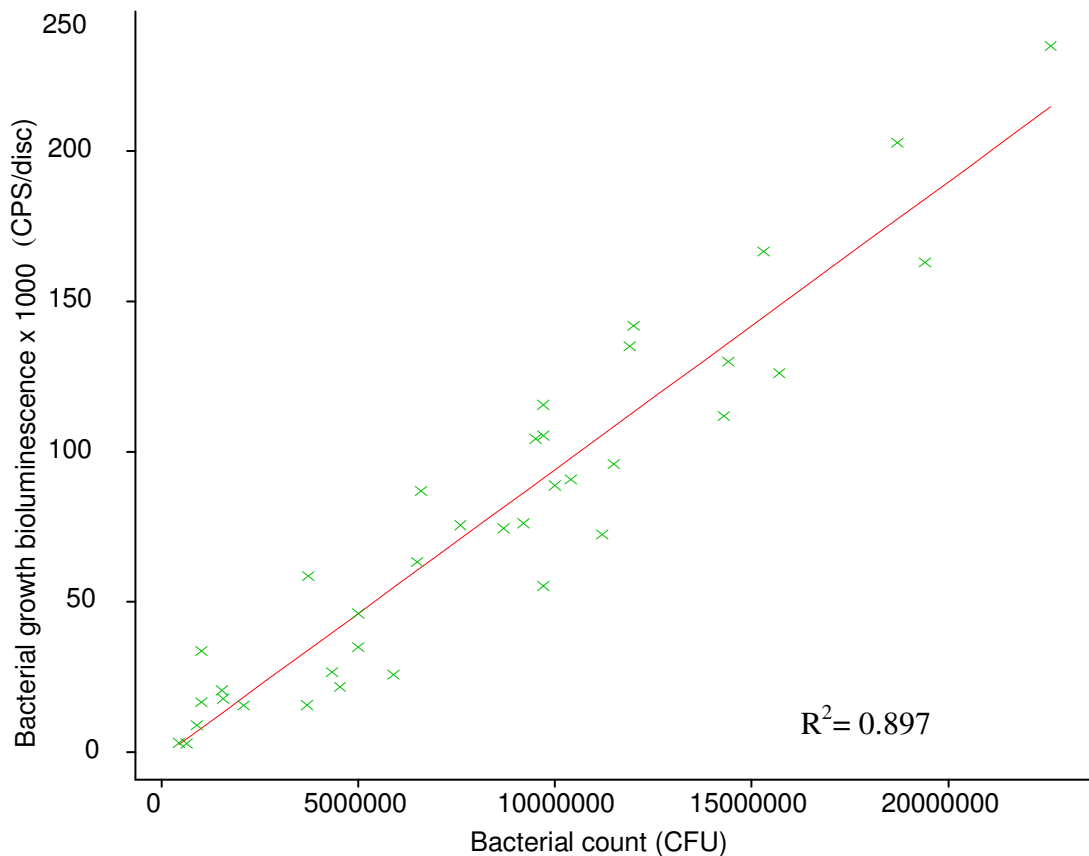


Figure 3.1 Directly proportional relationship between bacterial count data from different genotypes of the Sorbo x Gy-0 RIL population and luciferase photon emission. Each data point is the mean value derived from the average of 3 luciferase luminescence values per plant measured in counts per second (CPS) /disc values and 3 dilution plate count values measured in colony forming units per plant (CFU) making 9 values per data point.

3.3 Visualising the infection process

The emission of light from the Lux-tagged *Pst* offered the opportunity to visualise the infection post uniform inoculation to establish if any proliferation pattern could be discerned. Photon-counting cameras and standard camera images demonstrate the bacterial load difference and therefore, basal resistance difference between two selected genotypes and their respective Lux values two days post inoculation (Fig. 3.2). These images show a typical correlation between increased chlorotic symptoms and increased bacterial density, however this phenotype is not always observed in all genotypes and therefore cannot be relied on to indicate basal resistance level.

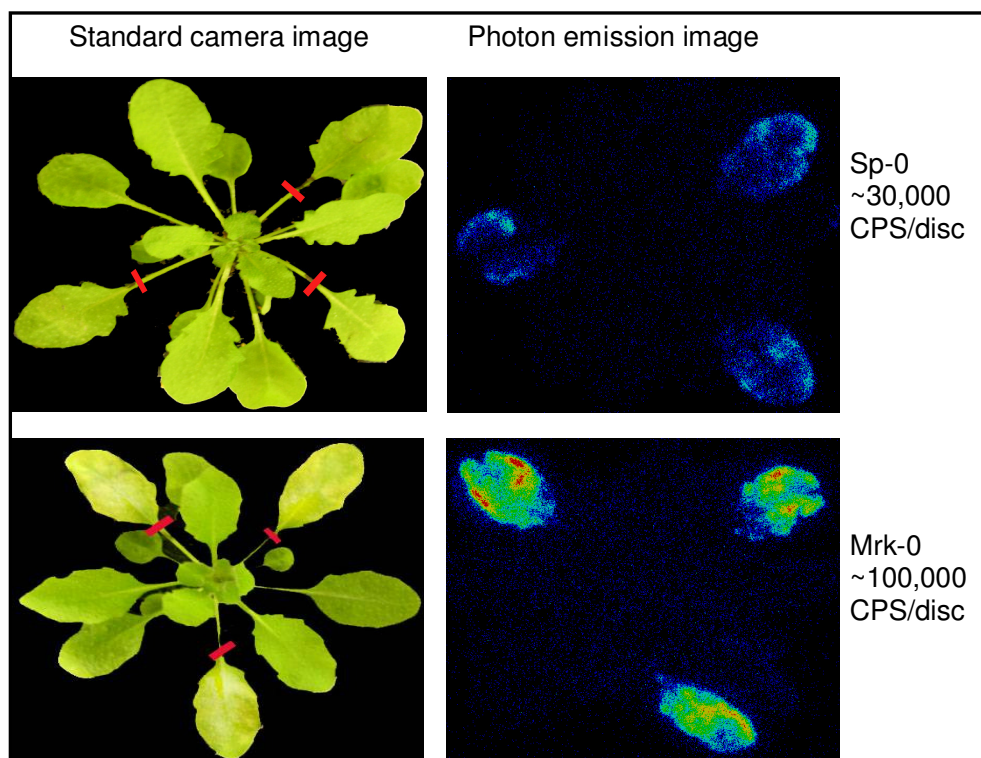


Figure 3.2 Comparative standard images and photon emission pictures taken 48 hrs after infiltration (OD_{600} 0.002) (exposure time 10 mins) for two genotypes and associated average basal resistance trait values. The red tabs indicate which leaves were infiltrated.

The photon emission images (Fig. 3.3) show a comparison between *Arabidopsis* genotypes which demonstrates the genotypic effect basal resistance can have on the proliferation of pathogenic bacteria. The images show similar patterns of high densities at the periphery of the leaves. The images also show that proliferation of *Pst* is limited to the infiltrated area for each of the shown ecotypes with significantly different basal resistance levels. The images show that even at high bacterial density, there is no *Pseudomonas* growth past the infiltrated area (Fig. 3.3, 3.4).

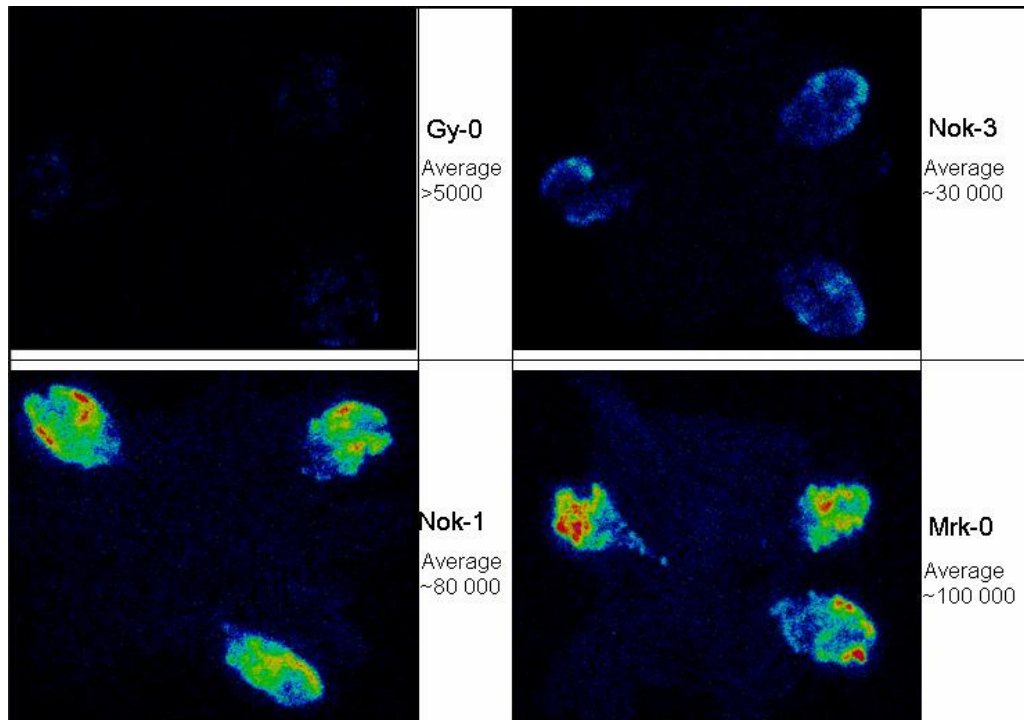


Figure 3.3 Light emissions from the infected leaf tissue over a ten minute exposure period 48 hrs post infiltration (OD_{600} 0.002) (exposure time 10 mins). Photon density is indicated by progressive heat map colour shading; blue= low photon density, Red= high photon density.

A selection of single frame images from the hourly photon emission (heat map) and luminescence (white) images were converted into a time lapse movie as displayed below (Figure 3.4). Sample images are at 24 hour intervals and show bacterial proliferation over the infection period which allows characterisation of proliferation. Typical proliferation is from leaf periphery spreading inwards to vascular tissue. There is no spread beyond the infiltrated area or utilisation of the vascular system to proliferate further than the infiltrated leaf. Comparison between images using plants subject to different infection methods indicate infiltrated leaves show a slightly more even distribution and faster growth than sprayed leaves. Proliferation rate may increase slightly during the first night cycle after infection, but after this there is no apparent circadian influence. However it should be noted that the plants are kept in a light-proof container throughout the time-lapse image recording, which may have an influence because sugar cycling and production will be adversely affected.

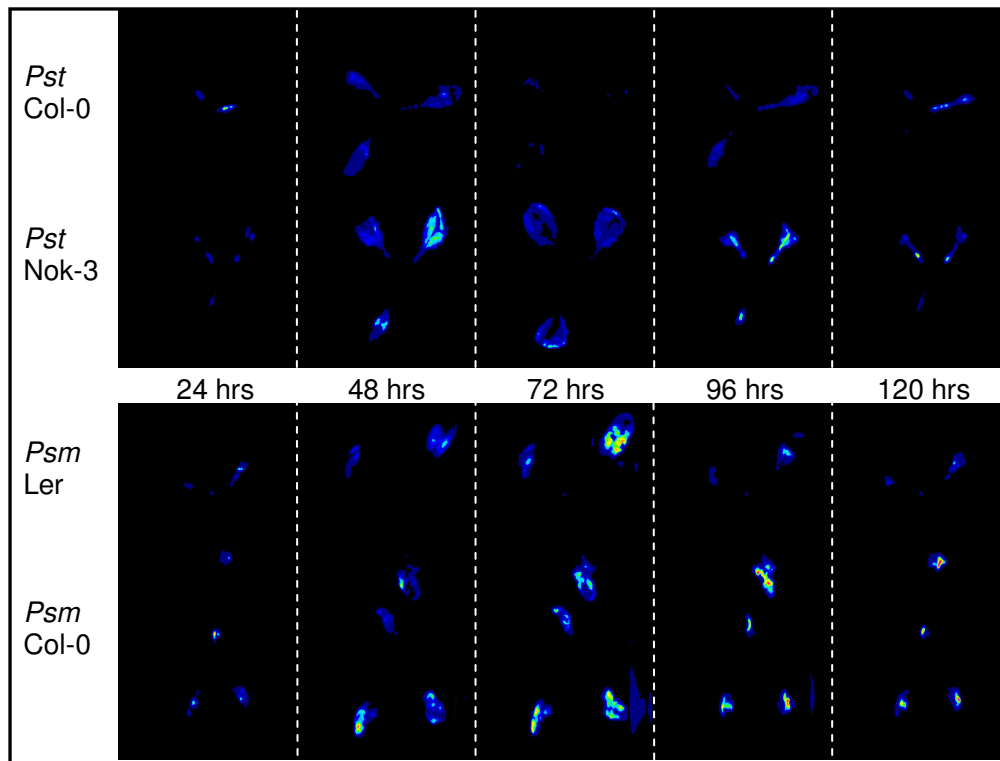


Figure 3.4 Photon emission images for Nok-3, Col-0 and Ler following infiltration with *Pst* and *Psm* (OD₆₀₀ 0.002) (exposure time 10 mins). Sample images are shown at 24 hr intervals with two individual plants per image.

3.4 Parental RIL line testing

Using the same assay system, the bacterial growth in 92 *Arabidopsis* accessions was investigated and the results are shown in Figure 3.5. The efficacy of basal resistance is shown within the variation of this measured population, as some ecotypes have reduced *Pst* growth to very close to zero (measurable with this assay). In contrast, some ecotypes show high bacterial densities within their tissue following similar *Pst* infiltration and this observed variation highlights natural differences in basal resistance amongst this selected population. ANOVA analysis proved trait similarity within a genotype ($P = >0.05$) from the minimum of 12 Lux bacteria data points per genotype and trait distinctiveness between RIL genotypes ($P = <0.05$). This trend was found consistently in RIL population testing (Fig. 4.2-4.6) which enhances confidence that the assay system can produce reproducibly similar data points from a RIL genotype and the assay is sensitive enough to distinguish between genotypes with varying basal resistance. As part of the ANOVA analysis F-test statistics are generated which defines the variation ratio around the mean when compared against

other variation from similar data sets. The F-test data suggests that the variance ratio between genotypes is significantly different ($F_{pr.} < 0.05$) as some genotype data sets show small deviation from the mean (e.g. less than 10%) and others show a large variation around the mean (e.g. over 60%). For accurate QTL analysis and for enhanced confidence in the assay system the variation ratios for RIL genotypes should be similar, therefore refinement of the assay system is required prior to QTL analysis. Natural genetic variation of basal resistance within the collection of ecotypes representative of the population is capable of reducing bacterial proliferation over 50 fold to almost no measurable bacterial growth (Fig. 3.5). Dissecting the components and mechanisms contributing to the observed variation is the aim of this study and requires a robust and reproducible assay system which will be refined in section 3.4.1.

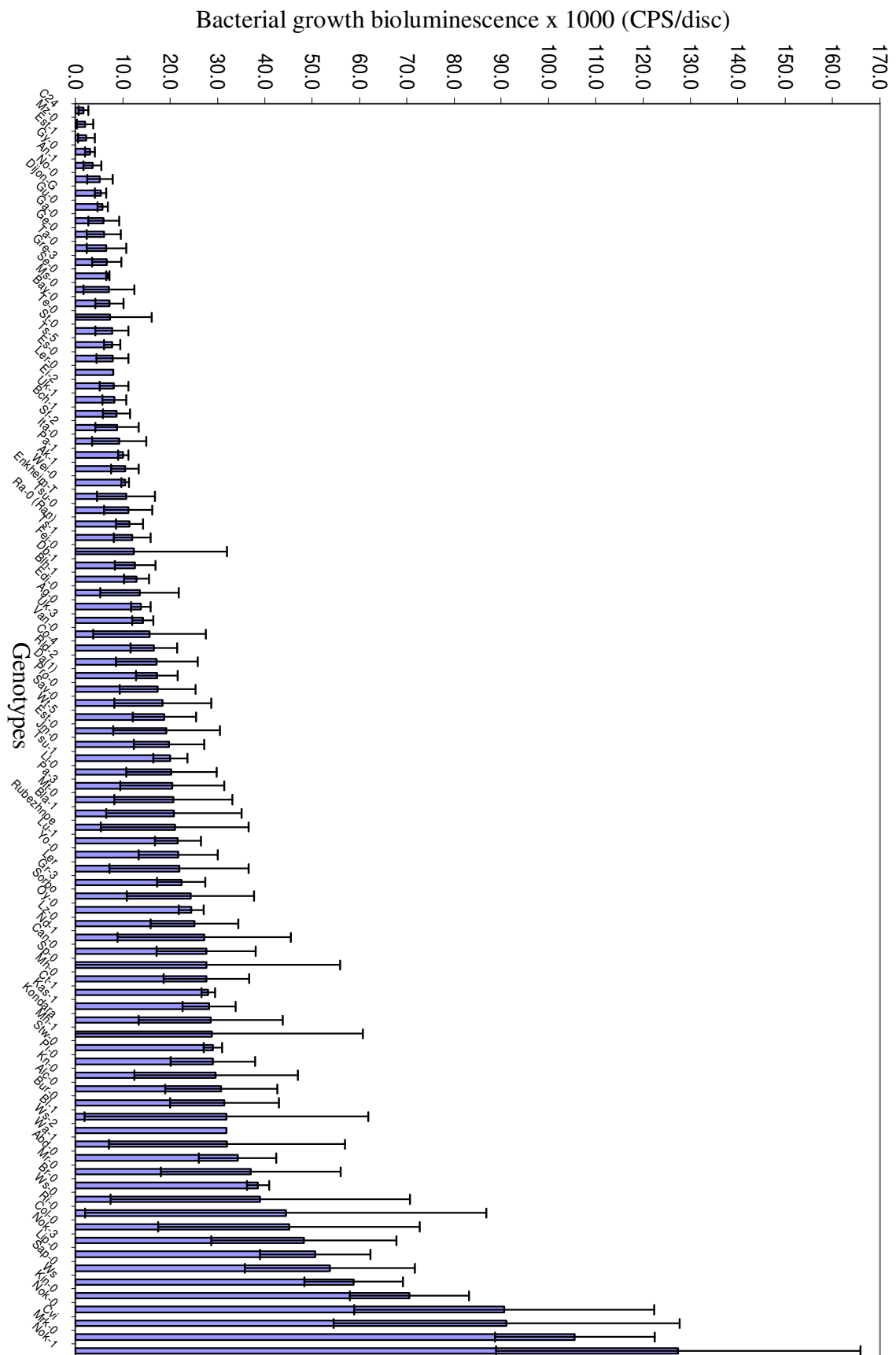


Figure 3.5 Range of basal disease resistance measured using the Lux assay produced a bacterial number value read in counts per second (CPS) /disc for 92 Parental RIL genotypes with four replicates per value (error bars indicate standard deviation).

3.4.1 Technique refinement

To establish a reproducible protocol for luminescence-based disease assay, which is suitable for further QTL analysis, a number of potential disease-related factors were investigated as summarized in Table 3.1. In addition, when RILs were analysed, internal controls were included into every plant tray to normalize calculations in order to reduce variation caused by any external or environmental factors. Significant differences were observed in the means and the variance of the Lux assay measure of bacterial growth when aspects of the experiments were modified, details of which are described in sections 3.4.2-.3.4.8. ANOVA tests were used to test the differences between means exposed to different treatments (P value) and F-test statistic (F pr.) were used to test if two or more population variances are equal by comparing the ratio of two variances away from the means, a value below 0.05 indicates significant variance.

Refinement experiment	Aspect of experiment modified to determine effect
Time-course of bacterial infection	Infiltration of plants and assessment of change in bacterial load over time e.g. 6,12,18 hrs after infiltration
Bacterial inoculum concentration	Bacterial infiltration concentration differences e.g. OD ₆₀₀ 0.002, 0.003, 0.004
Age of bacterial inoculum	Infiltration of plants with bacteria of different ages e.g. 6,12,18 hrs grown in KB medium
Plant age	Age of plant when infected e.g. 20,25,30 days
Time of day of inoculation	Infiltration of plants at different times of the day e.g. early, mid, late
Vacuum infiltration	Infiltration using vacuum infiltration instead of hand-infiltration
Other environmental factors	Tray position, growth room position, watering scheme etc.

Table 3.1 The experiments used to refine the luminescence-based disease assay method.

3.4.2 Bacterial time-course of infection

Three genotypes representing low, medium and high level of basal resistance were assessed over time for bacterial growth. The data shows that for genotypes Col-0 and Nok3, there was a clear increase in bacterial number over time, reaching a peak between 42 and 54 hrs, followed by a sharp decrease as plant tissue became damaged and desiccated. Gy-0 showed relatively little change over time as high basal resistance of this line reduced the bacterial content to relatively low levels throughout the screened period (Fig. 3.6). Following this experiment, the measurements were taken 48hrs post inoculation.

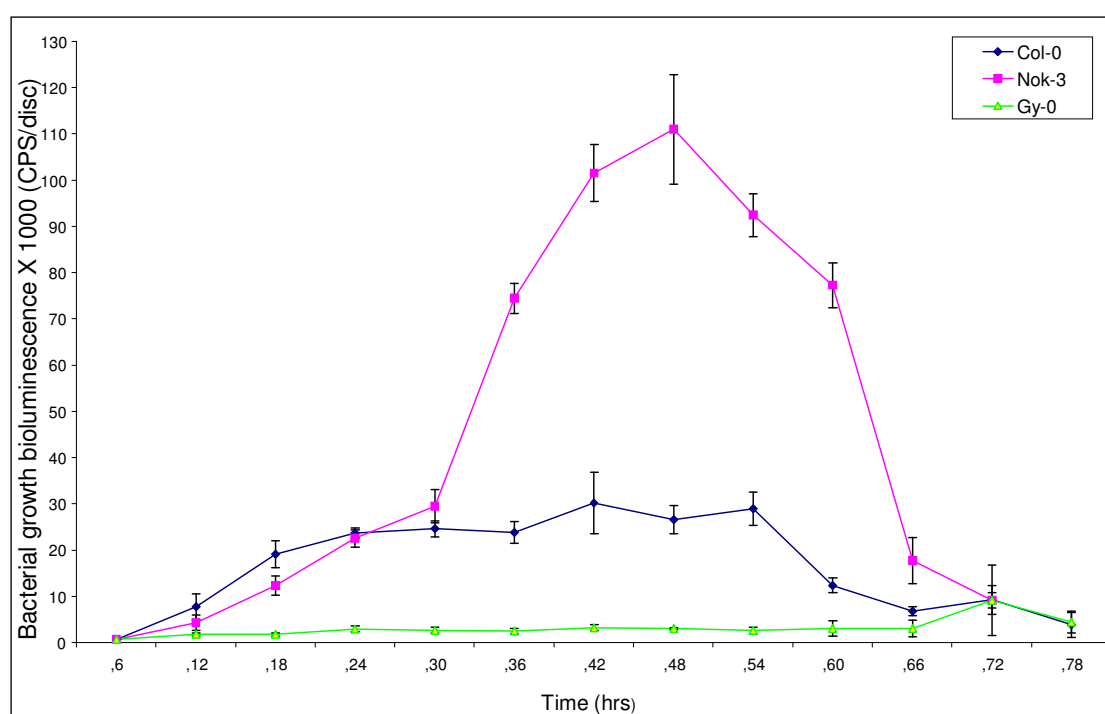


Figure 3.6 Time course of bacterial growth *in planta* in 3 genotypes. Plants were inoculated with *Pst* at OD₆₀₀ 0.002 and samples of 3 leaf discs per plant from 4 replicates were collected at 6 hr intervals to measure luminescence. The error bars represent standard deviation (SD).

3.4.3 Bacterial inoculum concentration

Col-0 plants were infiltrated (3 infiltrated leaves per plant and 4 replicate plants per treatment) with a range of bacterial densities and luminescence of inoculated leaves and were measured after 48 hrs. The results showed a relatively low Lux value for the highest density (OD₆₀₀ 0.004) of bacterial inoculum with the largest degree of

variation around the mean. Observations showed that at the tested time point a predominant surface area of this leaf tissue was severely damaged and desiccated for this infiltration concentration. The remaining three concentrations trialled showed the anticipated general trend of decreasing bacterial content as initial inoculum density is decreased. Mean bacterial growth is lowest at OD₆₀₀ 0.002 and ANOVA tests show its mean is statistically similar ($P = >0.05$) to the mean of OD₆₀₀ 0.003. Conversely, the mean of OD₆₀₀ 0.001 is significantly reduced ($P = <0.05$) in bacterial density due to the lower initial inoculum density compared to the OD₆₀₀ 0.002 and OD₆₀₀ 0.003 treatments (Fig.3.7). Variation around the mean is similar (F pr. >0.05) between OD₆₀₀ 0.001, OD₆₀₀ 0.002 and OD₆₀₀ 0.003. Variation around OD₆₀₀ 0.004 is higher than the other 3 treatments but variation is not significantly different (F pr. >0.05). Proportionally the OD₆₀₀ 0.002 has the lowest variability so when combined with its high mean value at the measurement point, the inoculum concentration was standardised to OD₆₀₀ 0.002.

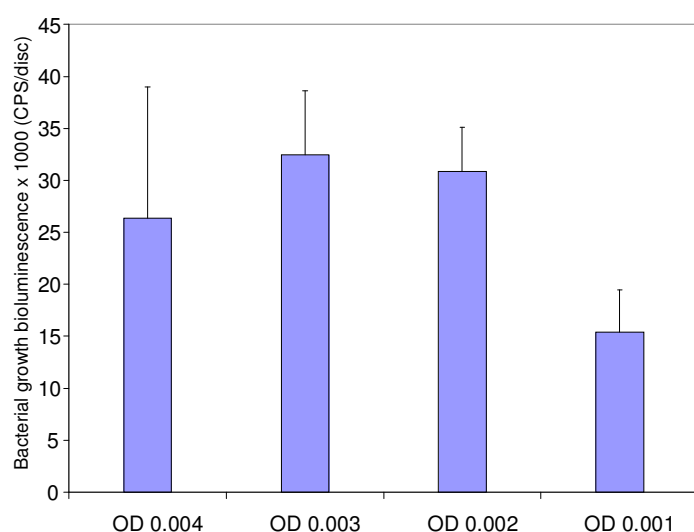


Figure 3.7 Effect of altering initial *Pst* inoculum density around the densities routinely used to assess bacterial growth. Col-0 ecotype tested 48 hrs after infiltration *in planta* from leaf discs. Three leaves from each of 4 plants were tested for each of the 4 selected inoculum densities. The error bars represent standard deviation (SD). ANOVA tests showed the means of OD₆₀₀ 0.004, 0.003 and 0.002 are statistically similar to each other however OD₆₀₀ 0.001 is significantly different ($P <0.05$) from OD₆₀₀ 0.002 and 0.003.

3.4.4 Age of bacterial inoculum

The bacterial incubation time prior to infiltration was varied to test bacterial growth differences. Col-0 plants were infiltrated with *Pst* (OD₆₀₀ 0.002) and luminescence was measured 48 hours later. ANOVA tests indicate the means from each of the incubation periods were not statistically different from one another ($P = >0.05$) however variation around the mean is of particular importance in improving assay accuracy. When comparing bacterial incubation at 12hrs and 36hrs, the variance ratio between these values is significantly different ($F_{pr} < 0.05$) however all other time points show no significant difference in variance. The earliest time point at 6 hrs shows the lowest bacterial proliferation 48 hrs post infection and a high but not significantly different degree of variability around the mean. The lowest variability around the mean was between 12-18 hrs of incubation time prior to infiltration with increasing variation from 24-36 hrs (Fig.3.8), suggesting that 12-18 hrs is the best age of inoculum to use.

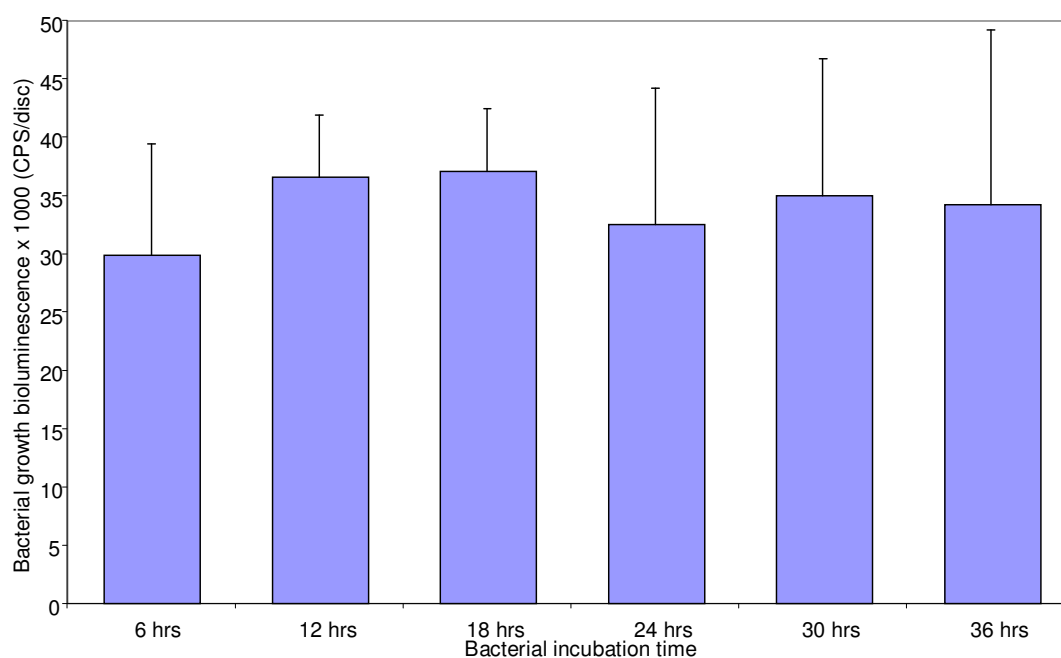


Figure 3.8 Effect on bacterial growth *in planta* of altering the age of the *Pst* inoculum. Bacterial inocula (OD₆₀₀ 0.002) were prepared from plates incubated for up to 36 hrs prior to infiltration into Col-0 leaf tissue. Bacterial growth was measured for 3 leaf discs per plant on 8 replicates plants per treatment. The error bars represent standard deviation (SD) no significant differences between bacterial incubation time means ($P = >0.05$).

3.4.5 Plant age

Six genotypes were chosen to represent the full range of basal disease resistance and infiltrated with OD₆₀₀ 0.002 of *Pst*. Starting at 20 days after seeds were sown, 4 plants of each genotype were infiltrated at 5 day intervals. The results show that for the majority of genotypes (Col-0, UK-3, Sorbo, Es-0), there was a slight increase in susceptibility with increasing plant age, typically around 35-45 days (Fig3.9). This coincided with many individual plants beginning to bolt however, two genotypes showed very little change around this period (Nok3 and Gy-0). This may suggest the potential for genotype dependent reallocation of resources at flowering time which may affect basal resistance. A study on the Vertifolia effect, a term describing the loss of agronomically costly partial resistance in a cultivar after several generations of selective breeding. During which, a major gene confers resistance to the dominant pathovar of the pathogen, has also shown a trade off between vegetative / reproductive growth and susceptibility of plants to pathogens (Van der Plank, 1963; J.Rant, pers. comm.). For each of the ecotypes both the means and associated variation were the most consistent compared with the variation observed across the time course at 30 days, this was selected as the measurement time point.

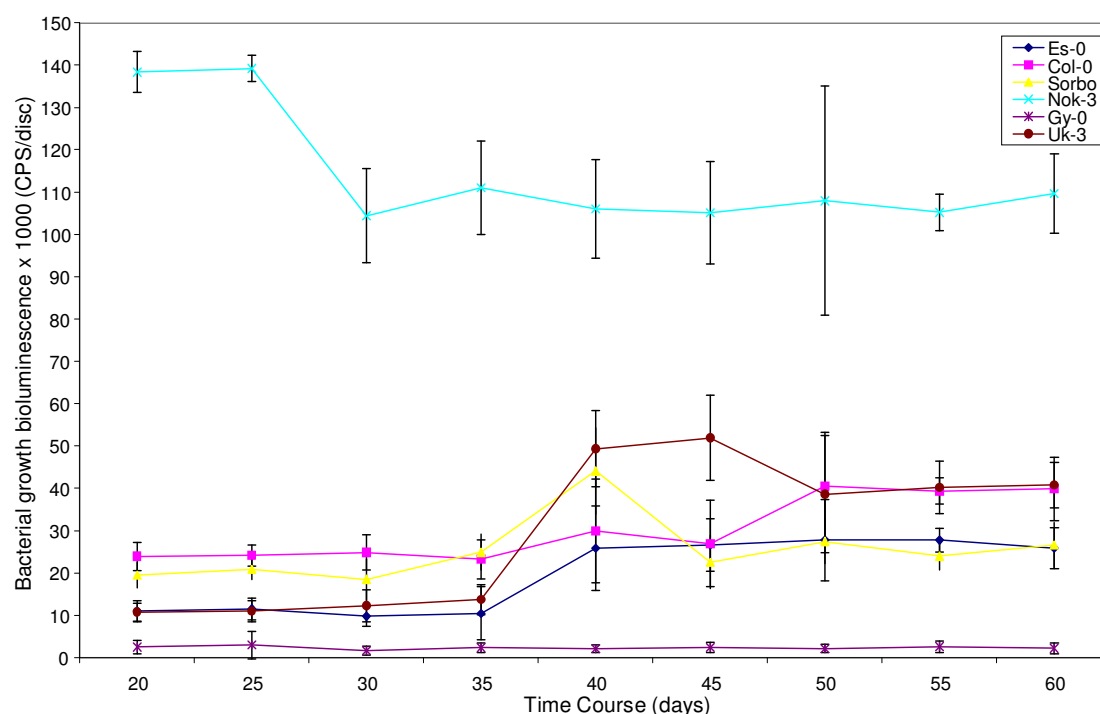


Figure 3.9 Impact of plant age on *Pst* growth. Six *Arabidopsis* genotypes inoculated with *Pst* at 5 day intervals. Three leaf discs in 4 replicate plants were scored for bacterial growth 48 hrs post infiltration (OD₆₀₀ 0.002). The error bars represent standard deviation (SD).

3.4.6 Time of day influence on bacterial growth

Bacteria were inoculated into Col-0 plants at different times of day to assess its impact on bacterial growth. Plants were infiltrated at 3 time points representing the start, middle and end of the light period and 3 leaf discs from each of 12 plants were assessed 48 hrs post inoculation. ANOVA test results show no significant difference ($P = >0.05$) between the means for each of these treatments. Variance ratio was statistically similar ($F_{pr} >0.05$) at its lowest level during the 1 pm time period, but not by a substantial difference (Fig. 3.10). From these results, it appears that there was no significant effect of time of day of inoculation on basal resistance. However, infiltration time was standardised into the same time period (10am-2pm) in case effects were subtle or varied in different ecotypes.

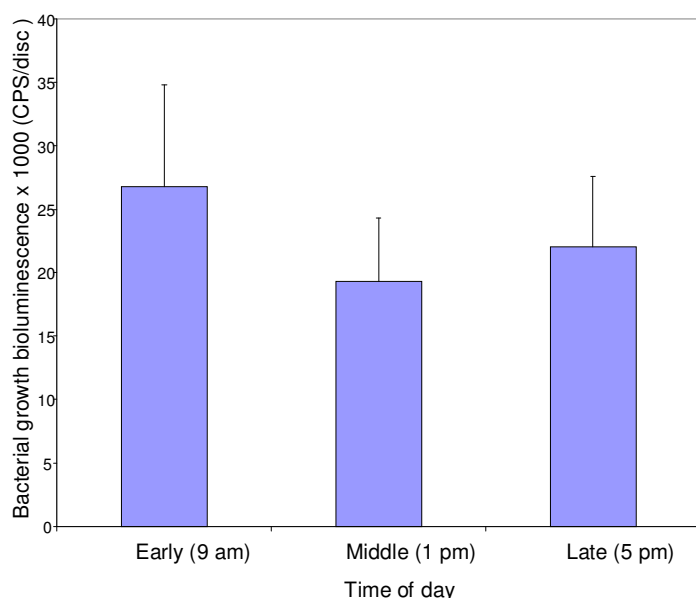


Figure 3.10 Effect on bacterial growth of inoculating Col-0 plants at different times during the light period. The light period was 09.00-18.00. Measurements were made 48 hrs after inoculation (OD_{600} 0.002). Twelve replicate plants were tested, taking 3 leaf discs per plant. The error bars represent standard deviation (SD). ANOVA tests showed no significant difference between Lux bacteria means of the different time periods ($P = >0.05$).

3.4.7 Infiltration method

There are several widely reported methods for introducing a bacterial pathogen into plant tissue however, to assess basal resistance in *Arabidopsis*, a uniform inoculation across an area of leaf tissue was required. A screen was conducted of 8 plants with 3 leaf discs taken per plant after exposure to an OD₆₀₀ 0.002 via 4 commonly used infection methods to determine which would be more effective. ANOVA test results show that vacuum infiltration is significantly ($P = <0.05$) less effective than dipping, spraying and pressure infiltration. Dipping is significantly ($P = <0.05$) less effective at the tested time point (48 hrs post inoculation) when it is not accompanied by a surfactant to lower water tension. At the tested time point, there is no significant difference ($P = >0.05$) in bacterial growth between plants infected using dip with surfactant, spraying and pressure infiltration. Variation around the mean as defined by the SD bars indicates that spraying and pressure infiltration shows decrease in variance ratio when compared to bacterial dipping and spray but this is not statistically significant ($F_{pr} >0.05$). Pressure infiltration over all shows the highest degree of reproducibility compared to the other methods and therefore will be used throughout the trials of basal resistance effect on *Arabidopsis*.

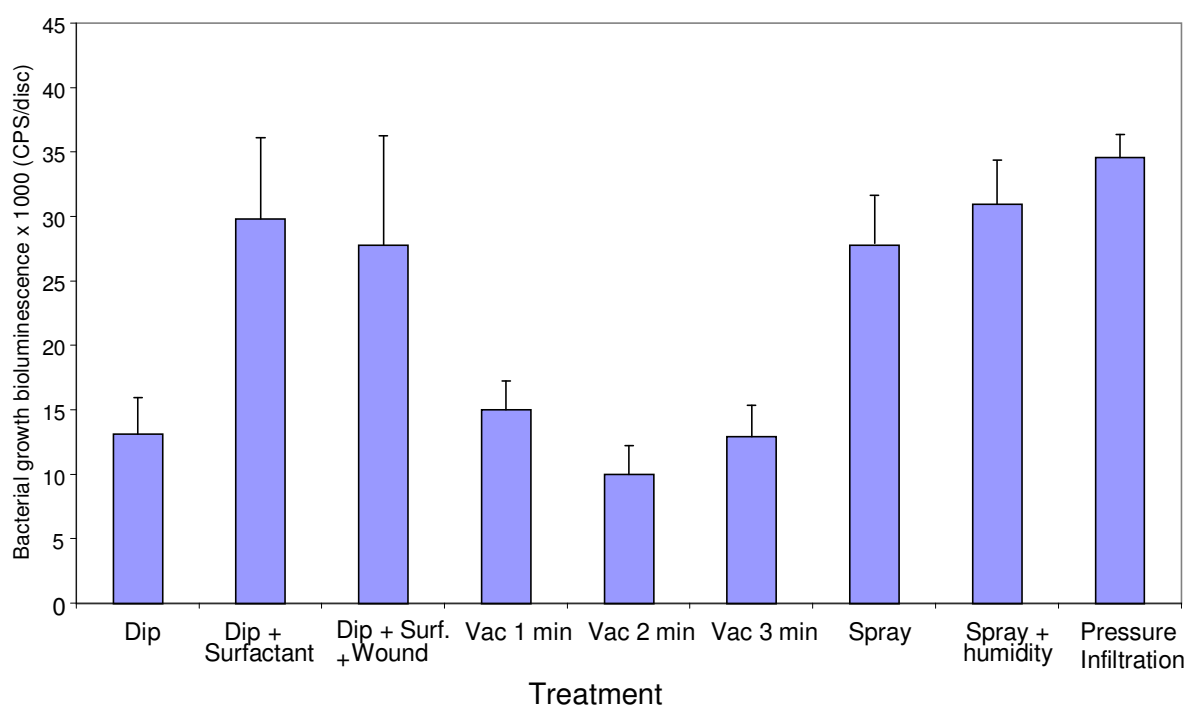


Figure 3.11 Effect of altering *Pst* infiltration method to compare dip, spray, vacuum and pressure infiltration to infect Col-0 plants. 8 plants per treatment were infected with an OD₆₀₀ 0.002 and three leaf discs per plant were assessed in the luminometer 48 hrs post inoculation. The error bars represent standard deviation (SD).

3.4.8 Environmental factors

To discount the possibility of position within the growth room having an effect on basal resistance due to, for example, differences in light intensity or temperature, a screen basal resistance values were measured for plants in different locations (Fig. 3.12). Three leaf discs were taken from 8 plants per treatment and infiltrated with an OD₆₀₀ 0.002: subsequent measurements were taken 48 hrs post inoculation. The ANOVA test results showed no significant ($P = >0.05$) differences in basal resistance response between trays grown in different areas of the growth room. The results show no discernable alteration to the variation around the mean (F pr. >0.05) for different positions within the growth room which allows for further large scale trials to occupy all areas of the growth room without influence on basal resistance.

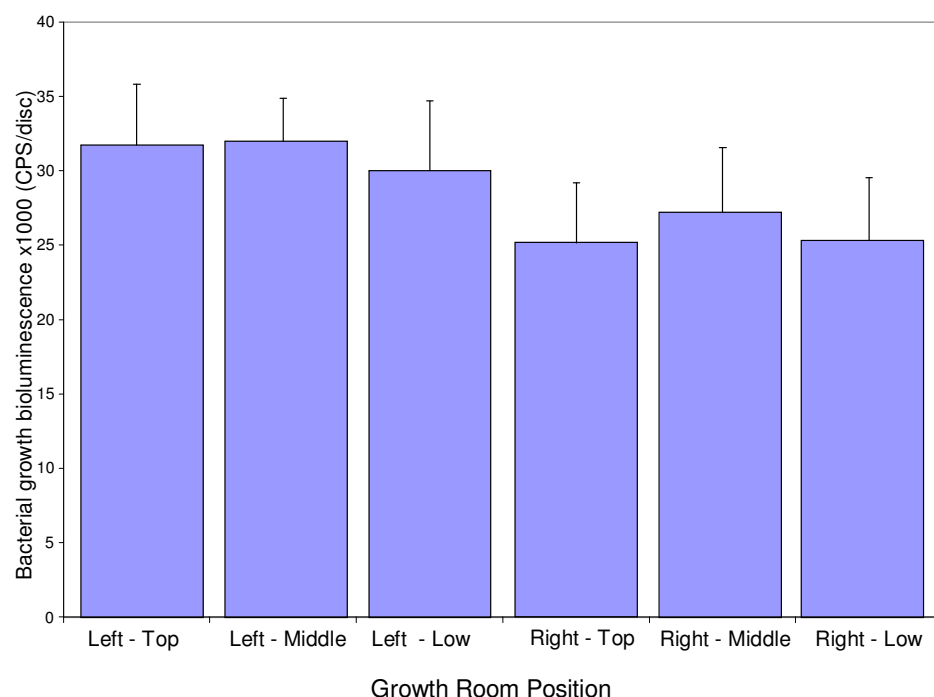


Figure 3.12 Effect on *Pst* growth after altering relative shelf positions for Col-0 trays positioned in the growth room. 8 plants were infiltrated with an OD₆₀₀ 0.002 and measurements were taken for 3 leaf discs per plant 48 hrs post inoculation. The error bars represent standard deviation (SD). ANOVA test results showed no significant difference between Lux bacteria number and growth room condition ($P = >0.05$).

The effect of a watering regime on basal resistance was assessed on a sample of 8 plants per treatment, where 3 leaf discs were taken per plant. A fixed volume of water (300ml) was supplied to each tray at different frequencies e.g. 2 days, 3 days etc. The effect of this was that trays supplied with the water on a daily basis had damp soil throughout their growth whereas trays provided with water only every five days had soil which was dry to the touch. The ANOVA test results showed that there was no significant difference ($P = >0.05$) between the watering treatments. Variation ratio around the mean was statistically similar (F pr. >0.05) for each of the treatments, however, variation did increase for those treatments at either end of the watering treatment range. Daily watering of a tray and the 5 day gap between watering treatment showed a higher variance ratio from around the mean (e.g. ~18% 5 day, ~13% 2 day) than treatments such as at 2 days where soil was moderately moist. These results indicate that watering regime should be standardised at 2 days to reduce variability (Fig. 3.13).

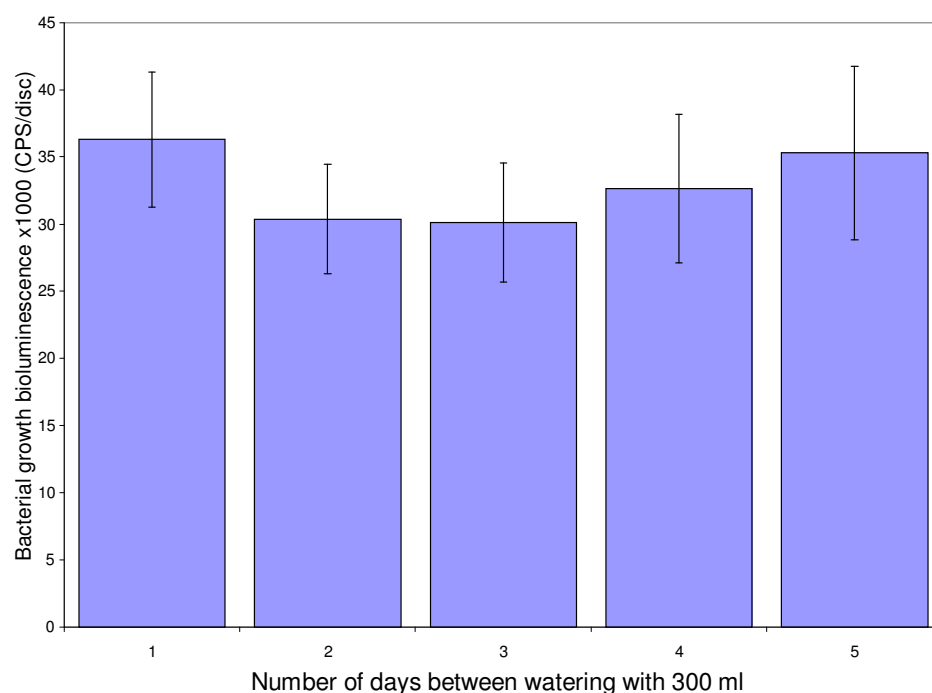


Figure 3.13 Effect on *Pst* growth of altering watering pattern for the Col-0 plants. Alterations increase sequentially the drought stress by increasing the time between watering the trays with a fixed water volume of 300ml. Plants infected with an OD_{600} 0.002 and measurements taken 48 hrs post inoculation. The error bars represent standard deviation (SD). ANOVA test results showed no significant difference in the means for watering treatments ($P = >0.05$).

3.5 Discussion

3.5.1 Natural variation in basal resistance of an *Arabidopsis* population

The results of Fig. 3.5 show a substantial range of natural variation of basal resistance across the *Arabidopsis* population, which was statistically distinct between ecotypes ($P = <0.05$). This assay shows that basal resistance alters the extent of *Pst* growth by more than 50-fold from practically zero bacteria to well over 100 CPS/disc, which equates to high bacterial densities within plant tissue. *Pst* was selected to be a universally virulent so basal mechanisms and action were highlighted and *R* gene action was not a contributing factor influencing resistance (Fan et al., 2008). A similar assay suggested that this variation in basal resistance affecting bacterial growth was in excess of 100-fold (Fan et al., 2008). The range of bacterial growth is continuous with no sharp steps in bacteria growth between genotypes which suggest that differences in basal resistance for closely scoring lines may be subtle genetic differences influencing, possibly influencing transcription rate or relative defence compound concentration. Mechanistic basal resistance differences between lines scoring widely apart (50-fold) may be more heavily influenced by substantial genetic sequence differences which impact on subsequent defence compound induction. The mechanisms and components underlying this range of variation which are capable of reducing bacterial growth to very close to zero are not clearly understood and are the objective for this project.

3.5.2 Refinement of technique

This study represents the first attempt at using the newly developed Lux assay on a wide array of genotypes and to generate trait data for detailed QTL analysis. Preliminary trials were conducted using the original protocols which were based on dilution plating procedures (Katagiri et al., 2002). A significantly high degree of variability (F pr. <0.05) was observed using these protocols to measure variation between RIL genotypes which would have greatly decreased the accuracy of any subsequent QTL analysis. Therefore, a series of refinement experiments were conducted to characterise the interaction and reduce variation.

Comparison between Lux values and bacterial number determined through plate assay showed a directly proportional relationship with an R^2 value of 0.897

(Fig.3.1). A significant positive correlation ($P = <0.05$) between Lux data and bacterial number confirms that Lux values accurately represent viable pathogen content within infiltrated leaf tissue. The fundamental principle of the assay is therefore sound, however, reproducibility must be increased for accurate QTL analysis (Fan et al., 2008). It should be noted red-spectrum photons are released from plant tissue as a result of the de-excitation of electrons within plant chlorophyll returning to ground state (Baker, 2008). A background reading was taken from uninfected tissue for each genotype and consistently found to be <500 RLUs. More sensitive photon monitors have used this fluorescence as a probe to study photosynthesis rates (Baker, 2008). In this study the contrast between bacterial luminosity shows background chlorophyll fluorescence to be an insignificant proportion (typically $<1\%$ of total Lux reading) of the majority of readings therefore can be discounted.

A time-course study of bacterial growth compared three genotypes over a 78-hour period (Fig. 3.6). All plants were infiltrated at the beginning of the experiment and selections of 4 plants per genotype were assessed for bacterial number at 6 hour intervals. This trial characterised the proliferation profile for each of the tested genotypes to show at what point after infiltration the bacterial population reached its maximum density and at which point variation was at its minimum. The trial demonstrated that in Nok3 and Col-0 the bacterial number peaked between 42 and 54 hrs after infiltration. Gy-0 showed only minimal differences between time points over the majority of the time-course as this line was selected to represent a high basal resistance variety. Peak bacterial density for Col-0 and Nok3 between 42 and 54 hrs showed a relatively stable period of trait variation therefore a time of 48 hrs post inoculation was selected as optimal time for trait scoring using the Lux assay. This conclusion was supported by the visualisation of infection time-course images as described by Fig. 3.4.

A variety of other experiments examined the variation produced when altering infiltration concentration (Fig.3.7), bacterial culture age (Fig.3.8), plant age (Fig.3.9) and time of day of infiltration (Fig.3.10). All of these experiments show some degree of variation of the means produced by altering the parameters, the majority of the trait results are not significantly different from each other, as only one aspect of the protocol has been modified. This was the predicted outcome from these experiments because these alterations aimed to characterise the interaction to determine which

factors were most influential. Modifications were made in and around the ranges of previously successful studies, as described in a variety of related literature, and therefore means were typically not drastically modified by the alterations. Examination of the variance ratios and standard deviation values was most relevant for these experiments because the aim of refinement was the reduction of variability and an increase in reproducibility between Lux values within and between plants and trays. For example, high bacterial concentration of OD₆₀₀ 0.004 produced a significantly higher variance ratio than the other infiltrated concentrations (F pr. <0.05). In other instances such as the number of days between watering the variance ratios were not significantly different however there was still a trend which suggested that trays watered every 2 days produced a lower variance around the mean than 5 days (~13% variance from the mean ~26% respectively). The protocol was subsequently revised with the majority of variables showing the least variability and most defined Lux values being selected.

In addition, bacterial growth was measured using different methods of inoculation e.g. spray and vacuum infiltration to assess the Lux means and variation around those means to select the most reproducible infection procedure for the assay. The comparison shows vacuum infiltration to produce Lux means significantly lower (P= <0.05) than both dip with surfactant, spray and pressure infiltration which were all statistically similar (P= >0.05). The variance ratios are similar (F pr. >0.05) between dip with surfactant, spray and pressure infiltration however pressure infiltration shows the highest mean and lowest variance ratio (~7%). Spray infiltration also has the complicating factor of genotypic variation in number of stomata, stomatal aperture, cuticle composition etc. which may affect bacterial colonisation of plant apoplast. Spray infection method treatment with flg22 was shown to produce a different expression pattern of defence-related genes compared with pressure infiltration of the same extract (Zipfel et al., 2004). Therefore pressure infiltration was selected to ensure a uniform infiltration of a consistent OD₆₀₀ 0.002 throughout the plant leaf for the assay system.

The factors most influential to the improvement of reproducibility included the reduction of the age of bacterial culture used to infiltrate (Fig.3.7), ensuring virulent bacteria had not been limited in growth through exhaustion of resources. OD₆₀₀ 0.001 produced significantly lower (P= <0.05) bacterial Lux measurement than OD₆₀₀ 0.002 and OD₆₀₀ 0.003. OD₆₀₀ 0.002 and OD₆₀₀ 0.003 treatments produced similar means

($P = >0.05$) and variance ratios (F Pr. >0.05) however OD_{600} 0.002 showed the smallest deviation from the mean. Standardising the infiltration concentration at OD_{600} 0.002 produces a steady rise in bacterial concentration within the leaf without rapid desiccation of tissue, as observed in the OD_{600} 0.004 infiltration, which showed a significantly higher variance ratio (F pr. <0.05) compared with OD_{600} 0.002. Bacterial proliferation rates and profiles were monitored for a selection of ecotypes (Fig. 3.6) and substantial variation in the mean Lux scores was observed. However, proliferation peaks for Col-0, Nok-3 and Gy-0 were significantly different from one another ($P = <0.05$) at 42 hrs to 54 hrs post inoculation yet within a genotype the means remained relatively stable for this time period, so measurements were taken at 48 hrs. This time may not be optimum for every genotype screened, but standardisation of measurement recording into this relatively consistent growth period should ensure a pragmatic compromise so the majority of genotypes are at the same disease stage at leaf disc collection.

In addition to protocol refinements, a review of the controls highlighted the advantages of the introduction of 4 Col-0 and 4 Ler controls spread throughout every p24 tray for more reliable QTL analysis. RIL plants and controls were randomised with the condition of at least one control on every column and row of a p24 tray which allows for a direct comparison with the controls for every RIL line. By comparing with internal controls, tray to tray variation and plant to plant variation was reduced by a minimum of approximately 50%. Protocol refinements aided reduction in the standard deviation around the mean so variance between genotypes is now of a similar ratio (F pr. >0.05). This continued to prove that the refined assay showed similarity within a genotype ($P = >0.05$) and accurately differentiated between RIL lines ($P = <0.05$). With accurate trait data combined with marker and genotype information, QTL analysis can be used to accurately define the chromosomal location influencing the trait.

3.5.3 Visualisation of infection

Following refinement of the Lux assay scoring technique, using sampled leaf discs from infiltrated tissue, a characterisation using qualitative observations of the proliferation pattern was conducted using the NightOwl photon emission camera.

After initial infiltration, a set of time-lapse photo images were generated that typically showed bacterial growth from leaf periphery spreading inwards to vascular tissue. There is little published data on proliferation patterns and there is no evidence of significantly different rates of resistance at the periphery or near the vasculature due to nutrient content or water content (Zwieniecki et al., 2003). This study proposes the hypothesis that this proliferation pattern is a result of gradually smaller intercellular space towards the leaf periphery. In a uniform inoculation of bacterial density over the leaf surface there will be a higher ratio of bacterial cells to plant cells at the leaf periphery. Therefore at the periphery, overall plant defence compound concentrations should be lower, as is the case with defensive alkaloids in Tobacco decreasing towards the leaf tip (Burton et al., 1992). The proliferation pattern is mirrored by the visual desiccation and chlorosis observed in other *Pseudomonas* interactions (Katagiri et al., 2002). Proliferation was limited to the infiltrated area, with no additional tissue spread or colonisation of the vasculature. *Arabidopsis* is not typically a host for *Pst* and therefore natural infection probably occurs on an opportunistic basis through wound sites. When introduced through infiltration, the pathogen is compatible enough to reproduce within initially infected areas, but not sufficiently adapted to overcome internal proliferation barriers (Katagiri et al., 2002).

Infection methods such as spray and pressure infiltration were compared, the result of which indicated that infiltrated leaves show slightly more even distribution and faster growth than sprayed leaves. This pattern was also observed in the comparison of infiltration techniques using the leaf disc Lux assay and is probably a result of the additional stomatal barrier that must be overcome. The sprayed bacteria are very reliant on the added surfactant to reduce surface tension around infection sites and are more sensitive to humidity variations than infiltrated leaves (Katagiri et al., 2002).

During the first night cycle after infection there appeared to be a slight increase in proliferation rate based on observation of the image sets, but after this there was no apparent circadian influence. However it should be noted that the plants used to take these measurements were kept in a light proof container for the duration of the 120 hrs over which the images were taken. Consequently it is likely that circadian rhythm, photosynthesis and respiration that are involved in sugar production would all have been adversely affected during this extended dark period. This may in turn influence bacterial proliferation rates and patterns compared with a plant in

typical 9/16 hrs daylight conditions. Refinement experiments taking readings at different points of the day did note some modest differences in the means at different point of the day, however these were not shown to be significantly different from one another. It may be worth repeating the experiment with several night time scoring points to assess any circadian influence. The images suggest that any difference would be subtle therefore this would need to be continued comprehensively in a further study focused on linking disease and metabolic pathway responses.

It was observed from the image sets generated that infiltrated leaves failed to search for a light source as other leaves did during this time, which could be evidence of a reduction / inhibition of phototropism. It is not clear from the images whether the age of the infected leaves caused the lack of phototrophic movement (as fully expanded leaves may stop tracking light sources) or whether a pathogenic inhibitory agent could be responsible. Further time-lapse studies could more effectively elucidate this possibility.

The use of photon emission cameras to characterise Lux tagged bacteria proliferation is a novel method which has not been previously used (as far as is currently documented). However, the results are only qualitative. A non-destructive quantitative assay could in future be created through the automated collection of software measurements of photon density within a given area (similar software already exists for automated cell counting from dilution plates, e.g. Invitrogen's Countless program). If images and readings could accurately track trait scores over the infection process to determine rate of change, then multiple QTL analysis could be conducted for the same RIL lines at different time periods. This could potentially determine if genes involved at the start of infection are same as those at later time points by monitoring any shifts in the QTL peaks between time intervals.

Chapter 4: QTL analysis of basal resistance with *Arabidopsis* recombinant inbred lines

4.1 Introduction

Quantitative trait analysis links a continuously-varying phenotypic trait with areas on a chromosome responsible for influencing that trait (Kearsey and Pooni, 1996). Quantitative trait loci (QTL) are chromosomal regions containing genes that contribute to a quantitative trait (Liu, 1998). Such quantitative traits may be encoded by several genes that are either involved in the same biochemical pathway or genes that may work in conjunction to produce an additive effect when all the responsible genes are expressed. To perform QTL analysis, a population of inbred lines which contain alleles that differentially affect trait expression is required, in addition to a set of genetic markers that characterise parental contribution to the inbred lines (Young, 1996; Koornneef et al., 1998). Quantitatively-measured trait data that varies continuously across the inbred population is another prerequisite for QTL analysis (Kearsey and Pooni, 1996).

An early successful application of QTL analysis came in *Arabidopsis* crosses (Landsberg x H51), used to dissect the underlying genetic components controlling flowering time, and identified 5 influential loci (Clarke et al., 1995). A QTL loci was coincident with *FRI*, a major late flowering gene, which in many early flowering accessions suffers loss of function. Another peak implicated the *RLN5* gene which was subsequently associated with affecting flowering time in short day length conditions (Clarke et al., 1995). From these initial associations, a family of *RLN(1-5)* genes was identified to be underlying the QTL loci, which conveyed significant influence over flowering time. It was subsequently shown that combinations of these early and late flowering alleles accounted for much of the wide variation range observed in the F2 segregating population (Clarke et al., 1995).

Following the identification of a QTL, characterisation of the genetic basis is achieved by assessing a set of individual candidate genes within the implicated region for their possible involvement (Weigel and Nordborg, 2005). Individual genes typically have a small phenotypic effect because of the complexity of interactions within the overall system. This makes it difficult to find definitive evidence for the identification of a QTL without supportive data (Asins, 2002). A number of

techniques have previously produced supportive evidence such as DNA polymorphisms that result in phenotypic effects and transgenic complementation. The knockout of a gene produces an effect on the trait due to its absence or significant differences in protein product functionality may also affect the trait (Weigel and Nordborg, 2005). Also, an allele implicated by QTL analysis may be re-introduced into a deficient plant line. This is realised through selective cross-breeding which screens the resulting progeny for restored function and so demonstrates a mechanistic link between the candidate gene and the trait (Weigel and Nordborg, 2005).

QTL analysis will aid in defining the number and location of the major components that contribute to basal resistance. From a survey of bacterial growth among 92 ecotypes previously used to generate the RILs, a subset of 5 RIL populations (comprised of at least 94 lines each) was selected to be further tested to produce quantitative trait data to act as a component in QTL analysis. Lux-tagged *Pst* assays of RIL lines consistently show a large (>50-fold) continuous range of variation of basal disease resistance that extends above and below the parental means across populations. QTL analysis produces a statistical indication of chromosomal location that underlies a trait. A reasonable range of error must be accounted for when selecting genetic markers to encompass a QTL peak, so potential candidate genes are not unnecessarily discounted. By cross referencing the genes within the QTL regions against these disease-related-gene lists, the candidate genes can be refined. The remaining candidate genes can be categorised based on their annotation information and compared with their alignment to QTL locations to determine if any correlation can be observed. From this analysis, a reduced and manageable set of candidate genes can be more closely examined and tested for their potential involvement in basal resistance and PAMP recognition.

4.2 Results

4.2.1 Selection of RIL lines

The graph below (Fig. 4.1) shows the same *Arabidopsis* genotype variation as described in Figure 3.5 with the addition of labels that illustrate the relative locations of the parents of the RIL populations chosen for subsequent QTL analysis. RILs were chosen that would be expected to occupy different ranges of variation primarily between the two parental values.

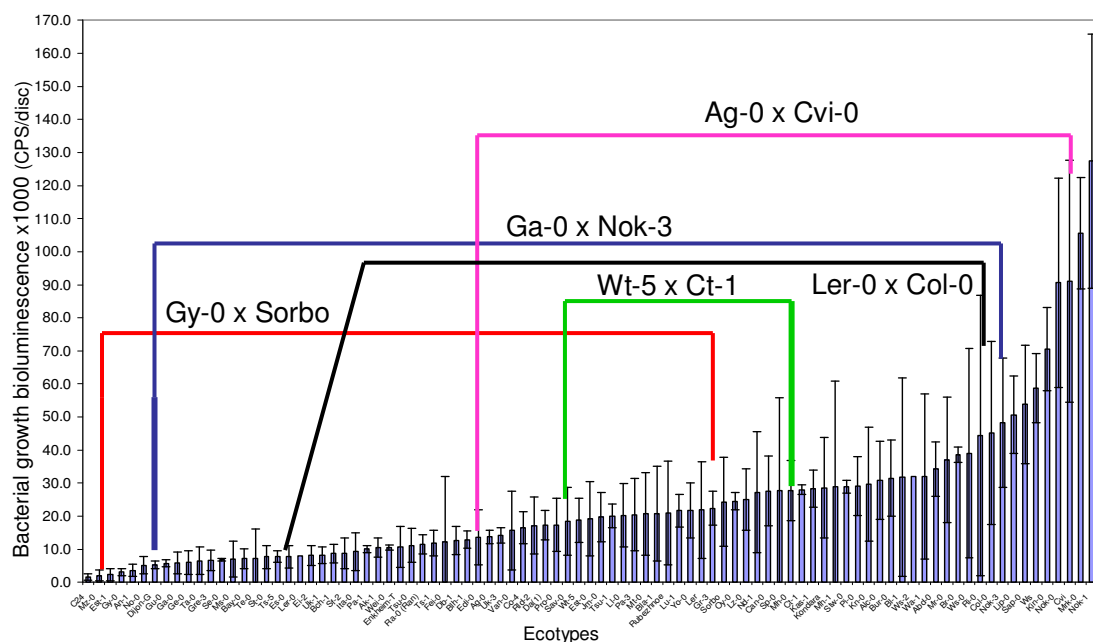


Figure 4.1 Selection of RIL populations for QTL analysis. *Pst* growth in 92 *Arabidopsis* accessions was determined previously and the parental accessions of RIL population were chosen as indicated to cover different ranges of variation of basal resistance. The error bars represent standard deviation (SD).

4.3.2 Basal resistance variation in recombinant inbred lines

The five sets of RILs derived from different parental crosses were specifically selected as they contained high marker density and had been extensively and accurately genotyped. Bacterial growth was tested to examine the range of variation over RIL populations for Sorbo x Gy-0, Nok-3 x Ga0, Cvi-0 x Ag-0, Col-0 x Ler and Wt-5 x Ct-1 (Fig. 4.2-4.6). RIL lines were selected to cover different ranges of basal resistance within the total population variation based on the preliminary ecotype screen and parental basal resistance values (Fig. 4.1). The subsequent segregating

RIL populations would primarily vary within those parental basal resistance boundaries. QTL analysis could help indicate whether variation at the highest susceptibility could be linked to the mechanisms resulting in variation at the lowest susceptibility by comparing analyses.

A consistent trend across the RIL populations shows a continuous range of variation which extends above and below the parental means. The ranges of the bacterial growth data across RIL genotypes do vary between RIL populations. RIL populations such as Nok-3 x Ga0 have a substantially larger range of bacterial growth than the RIL genotypes of the Col-0 x Ler population. Following assay refinement the genotypic variation of bacteria growth data around the mean for Sorbo x Gy-0 remains relatively low ($F_{pr} > 0.05$) and variance ratio remains statistically similar for the other surveyed genotypes (Fig. 4.2-4.6). Basal resistance effect on bacterial growth is genetically distinct as statistical analyses have shown significant consistency within a genotype ($P = > 0.05$) and distinctiveness between genotypes ($P = < 0.05$).

4.3.3 Trait data from Sorbo x Gy-0 RIL population

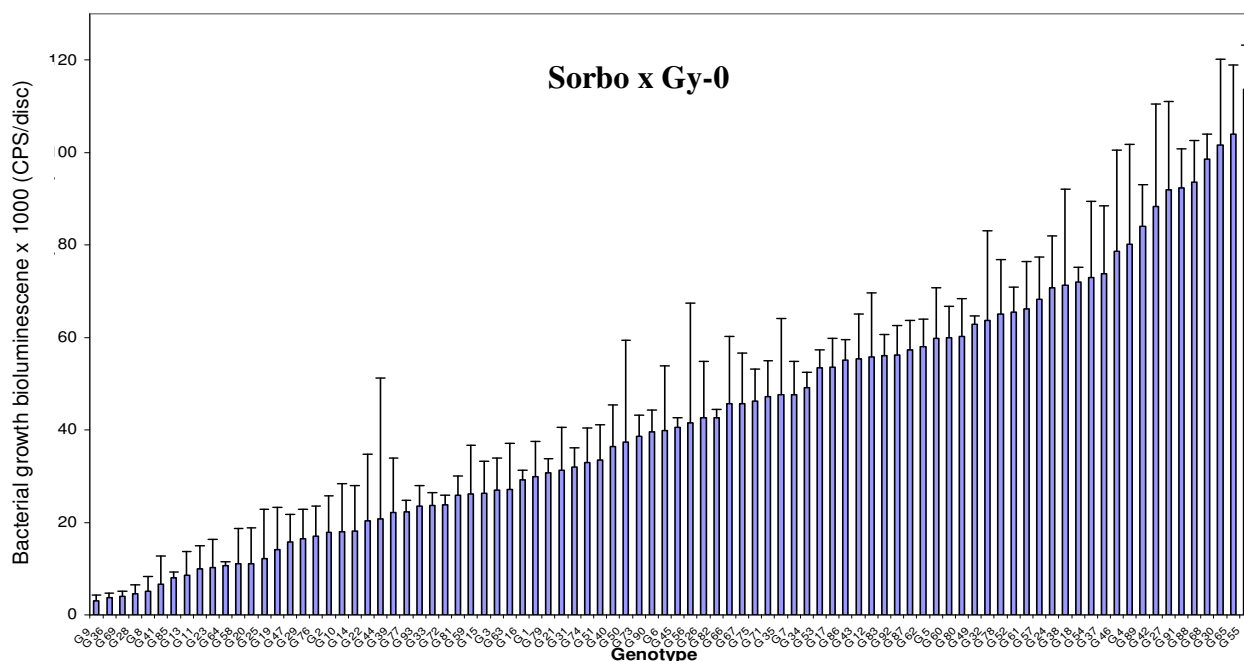


Figure 4.2 Bacterial growth trials for *Pst* infected leaf tissue showing basal resistance trait variation across the Sorbo x Gy-0 RIL population. An OD₆₀₀ of 0.002 was pressure infiltrated into 3 leaves per plants with at least 4 replicate plants per RIL. Bacterial growth was measured 48 hrs post inoculation. The error bars represent standard deviation (SD).

4.3.4 Trait data from Nok3 x Ga0 RIL population

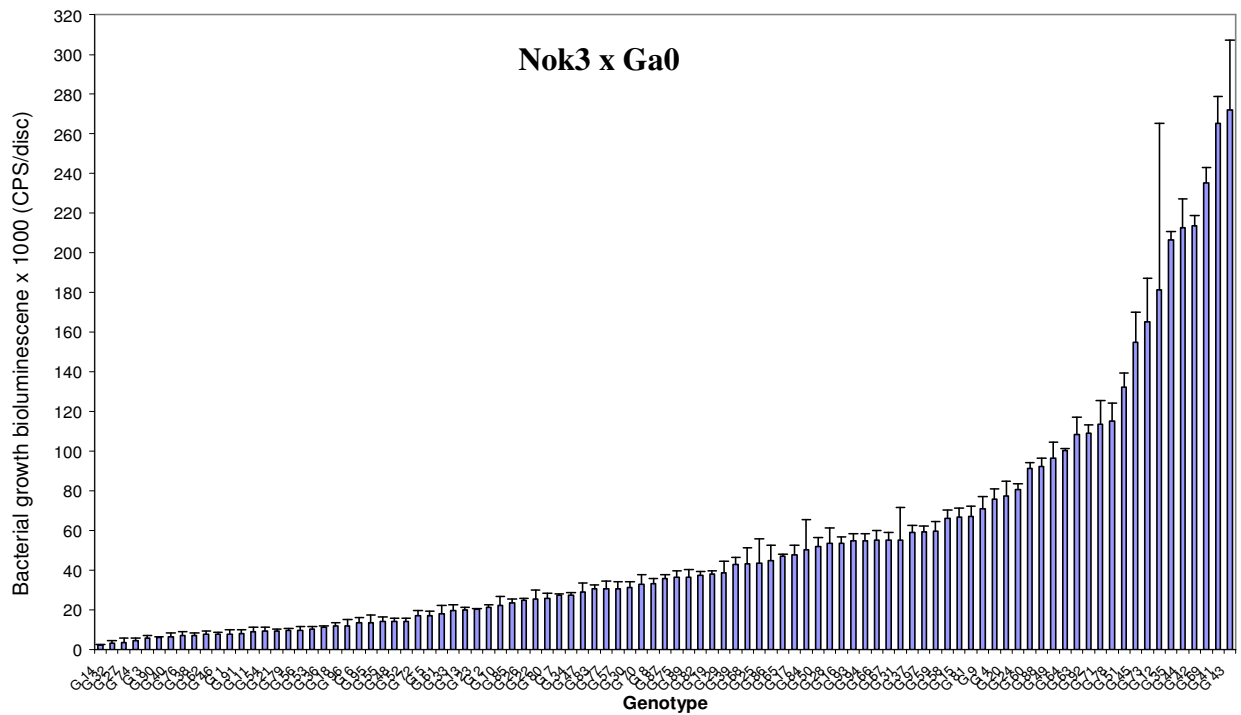


Figure 4.3 Bacterial growth trials for *Pst* infected leaf tissue showing basal resistance trait variation across the Nok3 x Ga0 RIL population. An OD₆₀₀ of 0.002 was pressure infiltrated into 3 leaves per plants with at least 4 replicate plants per RIL. Bacterial growth was measured 48 hrs post inoculation. The error bars represent standard deviation (SD).

4.3.5 Trait data from Cvi0 x Ag0 RIL population

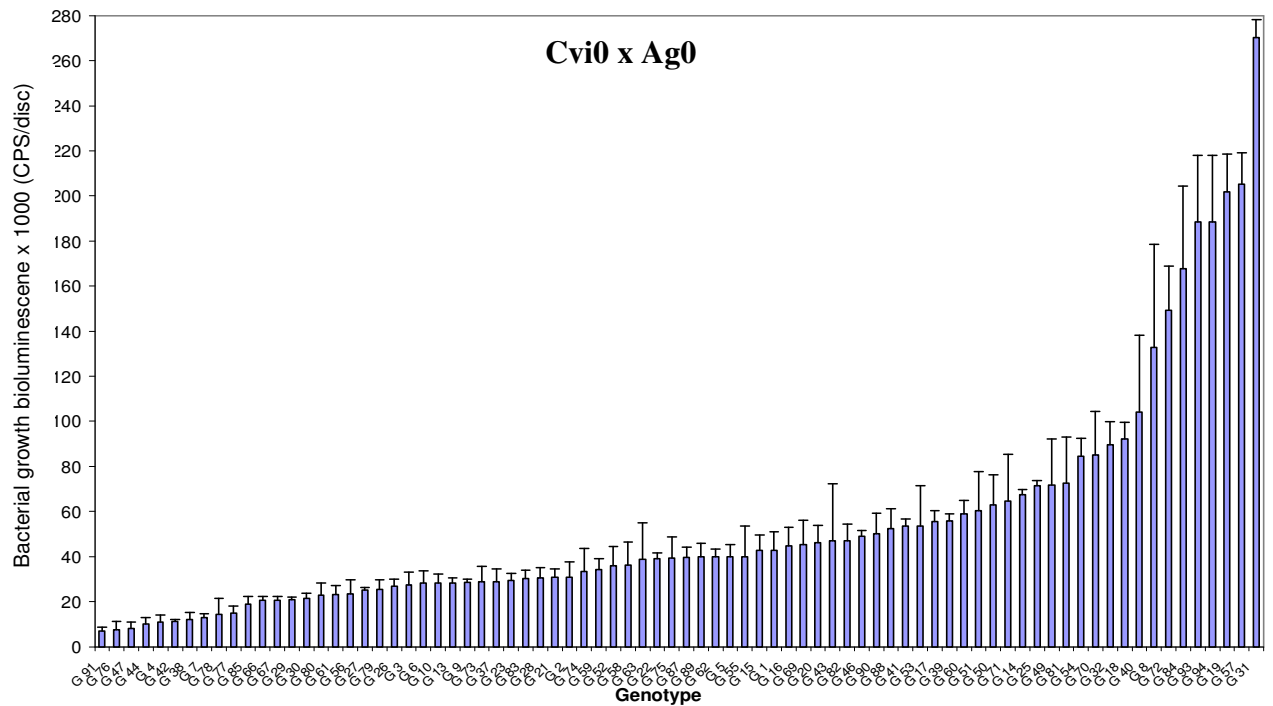


Figure 4.4 Bacterial growth trials for *Pst* infected leaf tissue showing basal resistance trait variation across the Cvi0 x Ag0 RIL population. An OD₆₀₀ of 0.002 was pressure infiltrated into 3 leaves per plants with at least 4 replicate plants per RIL. Bacterial growth was measured 48 hrs post inoculation. The error bars represent standard deviation (SD).

4.3.6 Trait data from Col-0 x Ler RIL population

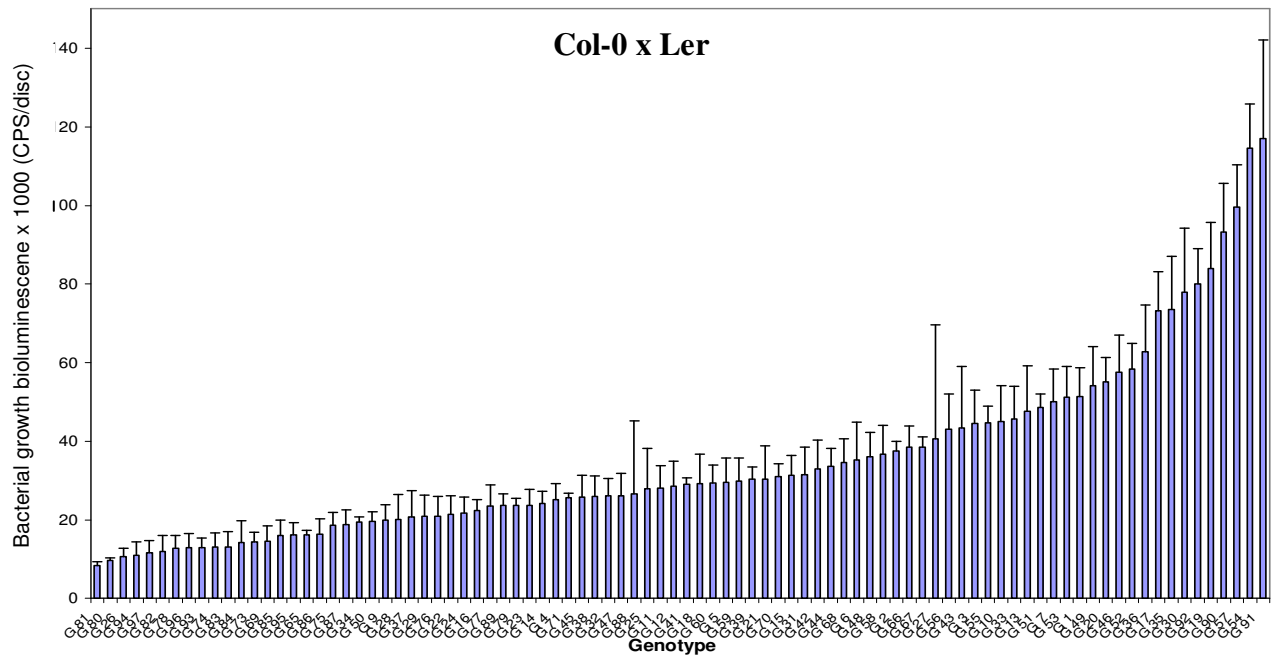


Figure 4.5 Bacterial growth trials for *Pst* infected leaf tissue showing basal resistance trait variation across the Col-0 x Ler RIL population. An OD₆₀₀ of 0.002 was pressure infiltrated into 3 leaves per plants with at least 4 replicate plants per RIL. Bacterial growth was measured 48 hrs post inoculation. The error bars represent standard deviation (SD).

4.3.7 Trait data from Wt-5 x Ct-1 RIL population

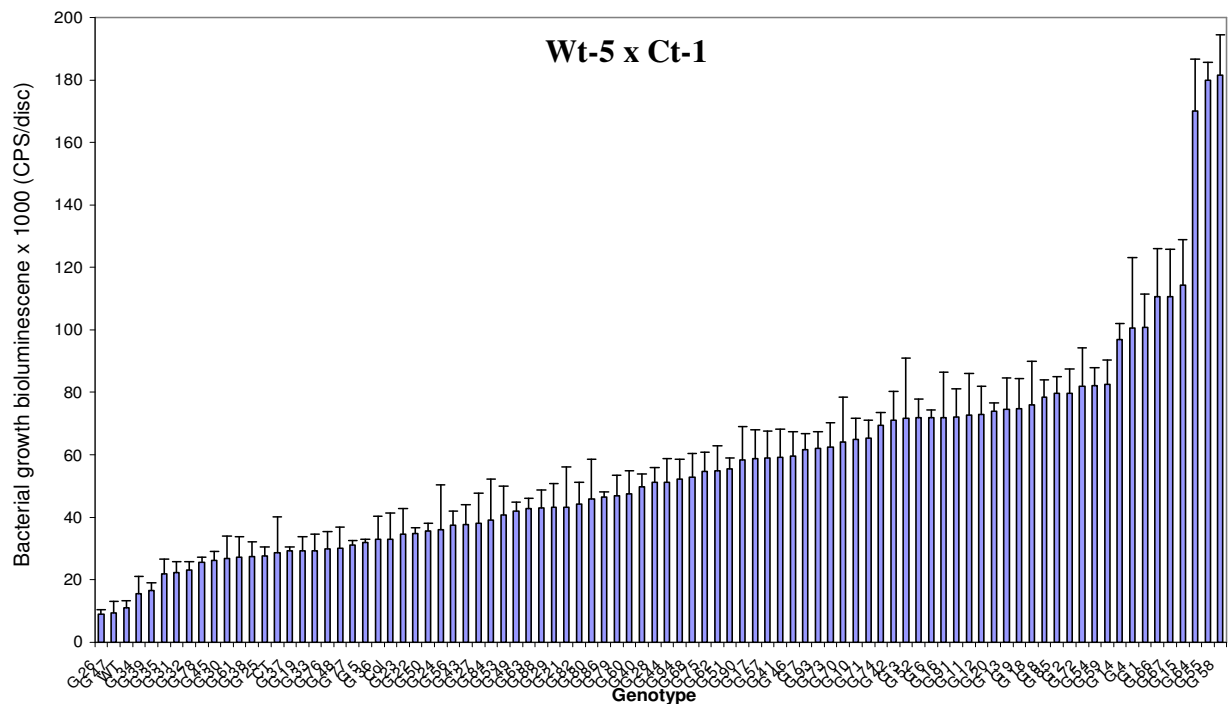


Figure 4.6 Bacterial growth trials for *Pst* infected leaf tissue showing basal resistance trait variation across the Wt-5 x Ct-1 RIL population. An OD₆₀₀ of 0.002 was pressure infiltrated into 3 leaves per plants with at least 4 replicate plants per RIL. Bacterial growth was measured 48 hrs post inoculation. The error bars represent standard deviation (SD).

4.4 QTL analysis

Independent QTL MQM graphs and analyses for each of the 5 RIL populations: Sorbo x Gy-0, Nok-3 x Ga0, Cvi-0 x Ag-0, Col-0 x Ler and Wt-5 x Ct-1 that were infiltrated with *Pst* are shown (Fig.4.7-4.11). Below, each graph has a table to compare LOD values with significant thresholds for the different chromosomes which may have multiple significant peaks for each chromosome. The table also shows the percentage of the variance explained by a peak in addition to additive effect scores. A positive additive effect indicates that the allele increasing disease was derived from the most susceptible parental line and a negative effect would show that the allele increasing disease was derived from the more resistant parent. Other analyses such as interval mapping and Kruskal-Wallis were conducted for each of these RIL populations. All analyses showed similar results and indicated the presence of QTLs in the locations described therefore the graphs below display the MQM mapping format, as this is the most rigorous and accurate analysis. A summary of relevant Kruskal-Wallis analysis is described below in the Table (4.1) to illustrate the correlation between analysis methods.

RIL line	Chromosome	Position (cM)	Marker	K*	Signif.
Sorbo x Gy0	1	66.00	*nga128	7.675	**
Sorbo x Gy0	3	13.00	*msd212	2.744	**
Sorbo x Gy0	5	44.00	*jc5136	11.409	*****
Nok3 x Ga0	1	13.00	*f16j7	4.817	**
Nok3 x Ga0	2	45.00	*nga112	4.963	**
Nok3 x Ga0	3	36.00	*k11j15	3.128	**
Nok3 x Ga0	4	6.00	*t19j18	22.481	*****
Cvi0 x Ag0	3	1.200	*nga162	3.133	**
Col-0 x Ler	3	4.12	*SGCSNP379	5.959	**
Col-0 x Ler	5	122.69	*agp11e	5.164	**
Wt5 x Ct1	2	53.00	*t32f646516	19.025	*****
Wt5 x Ct1	5	8.00	*nga150	3.463	**

Table 4.1 First analysis stage: summary of the of Kruskal-Wallis analysis data for significant QTL peaks. Significance value < *0.1 **0.05 ***0.01 ****0.005 *****0.001.

The ‘MapQTL’ software interpretation of the gradient towards the locus of greatest trait influence is a set of line graphs covering the length of each chromosome examined. The line measures the probability of a locus on influencing the trait, shown as a logarithm of odds (LOD) score. An ideal QTL peak has a high LOD score over a tightly-defined chromosomal area.

4.4.1 Sorbo x Gy-0

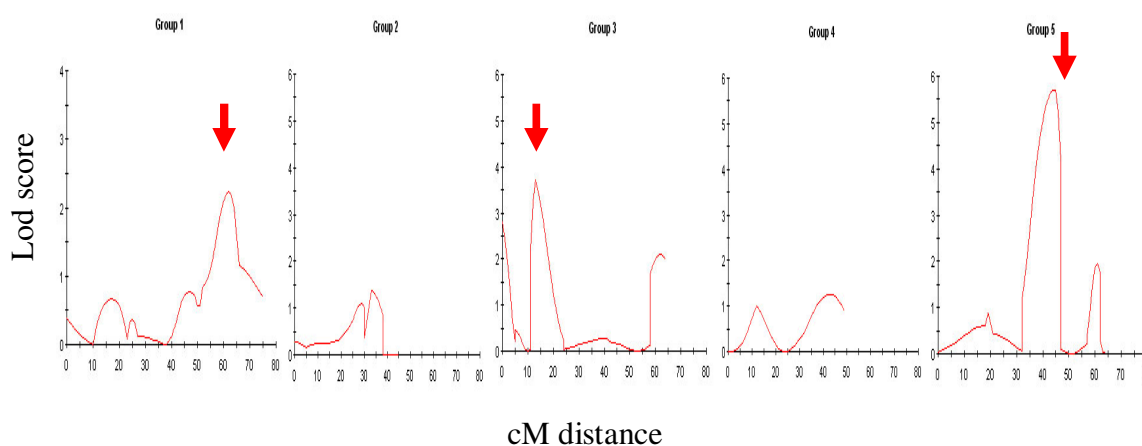


Figure 4.7 QTL linkage group analysis associated with RIL population from Sorbo x Gy-0 cross. MQM analysis using ‘MapQTL’ software displays peaks at chromosomal locations where trait variance is explained. Significant peaks on groups 1, 3 and 5 highlighted by arrowheads.

Chromosome	LOD score	LOD 95% sig. threshold	% variance explained	Sig.QTL	Additive effect
Group 1	2.31	2.2	31.7	Yes	+ 220.0
Group 2	1.39	2.3	7.8	No	N/A
Group 3	3.73	2.3	15.6	Yes	+ 33.2
Group 4	1.25	2.2	9.8	No	N/A
Group 5	5.71	2.2	32.3	Yes	+ 32.6

Table 4.2 Summary of MQM analysis showing Logarithm of odds (LOD) scores and confidence interval with % explanation of variance by QTL location. A positive additive effect indicates that the allele increasing disease was derived from Sorbo and a negative effect would show that the allele increasing disease was derived from Gy-0. Significant QTLs above 95% statistical confidence level. Significant peaks on groups 1, 3 and 5.

4.4.2 Nok3 x Ga0

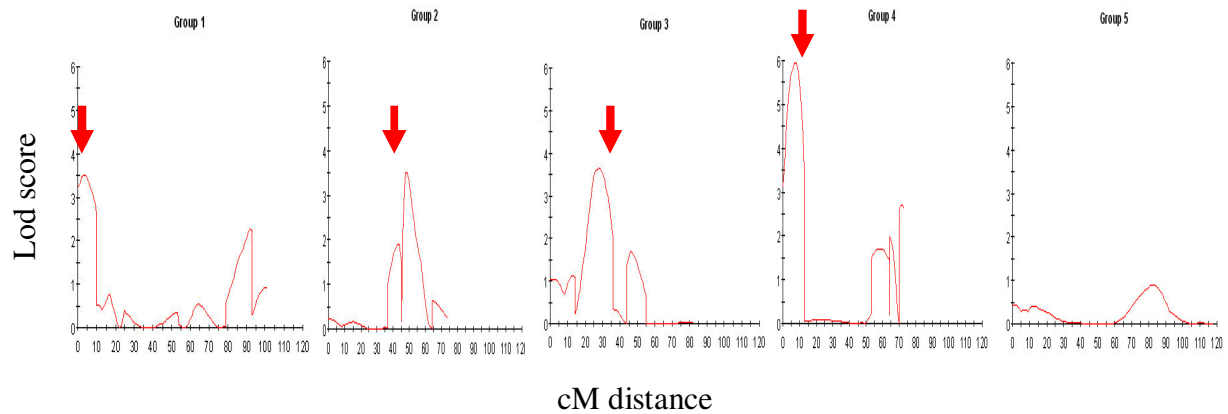


Figure 4.8 QTL linkage group analysis associated with RIL population from Nok3 x Ga0 cross. MQM analysis using ‘MapQTL’ software displays peaks at chromosomal locations where trait variance is explained. Significant peaks on groups 1, 2, 3 and 4 highlighted by arrowheads.

Chromosome	LOD score	LOD 95% sig. threshold	% variance explained	Sig. QTL	Additive effect
Group 1	3.51	1.9	6.6	Yes	-21.0
Group 2	3.52	1.7	6.6	Yes	+51.2
Group 3	3.64	1.6	12.4	Yes	+24.4
Group 4	5.94	1.7	14.6	Yes	+27.1
Group 5	0.89	1.7	4.6	No	N/A

Table 4.3 Summary of MQM analysis showing LOD scores and confidence interval with % explanation of variance by QTL location. A positive additive effect indicates that the allele increasing disease was derived from Nok-3 and a negative effect shows that the allele increasing disease was derived from Ga0. Significant QTLs above 95% statistical confidence level. Significant peaks on groups 1, 2, 3 and 4.

4.4.3 Cvi0 x Ag0

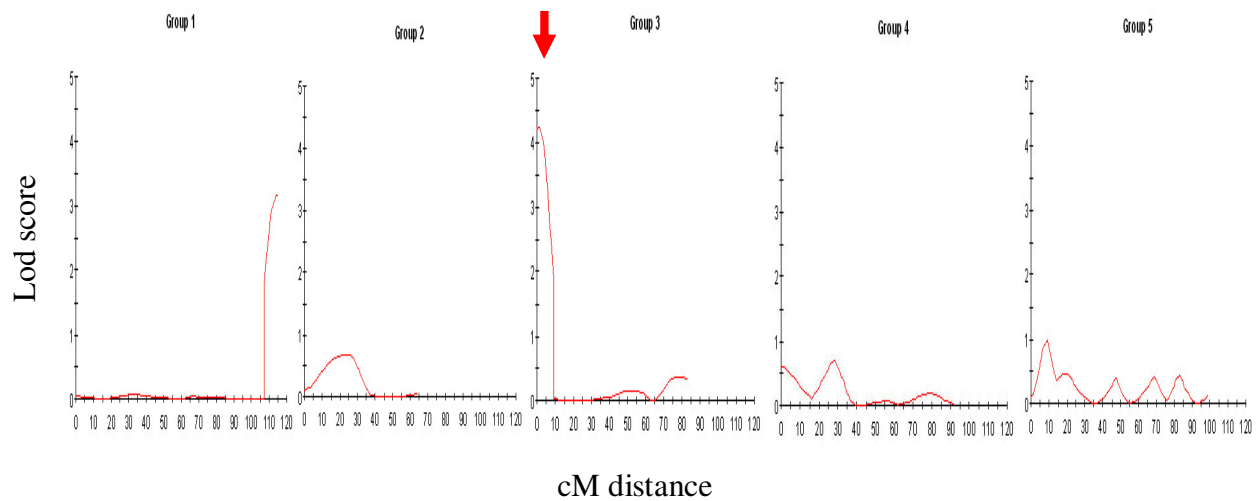


Figure 4.9 QTL linkage group analysis associated with RIL population from Cvi0 x Ag0 cross. MQM analysis using ‘MapQTL’ software displays peaks at chromosomal locations where trait variance is explained. Significant peak on group 3 highlighted by arrowhead.

Chromosome	LOD score	LOD 95% sig. threshold	% variance explained	Sig.QTL	Additive effect
Group 1	3.17	4.5	14.0	No	N/A
Group 2	0.69	1.6	2.9	No	N/A
Group 3	4.23	1.7	19.9	Yes	+26.3
Group 4	0.71	1.7	2.9	No	N/A
Group 5	0.99	2.0	3.9	No	N/A

Table 4.4 Summary of MQM analysis showing LOD scores and confidence interval with % explanation of variance by QTL location. A positive additive effect indicates that the allele increasing disease was derived from Cvi-0 and a negative effect would show that the allele increasing disease was derived from Ag0. Significant QTLs above 95% statistical confidence level. Significant peak on group 3.

4.4.4 Wt-5 x Ct-1

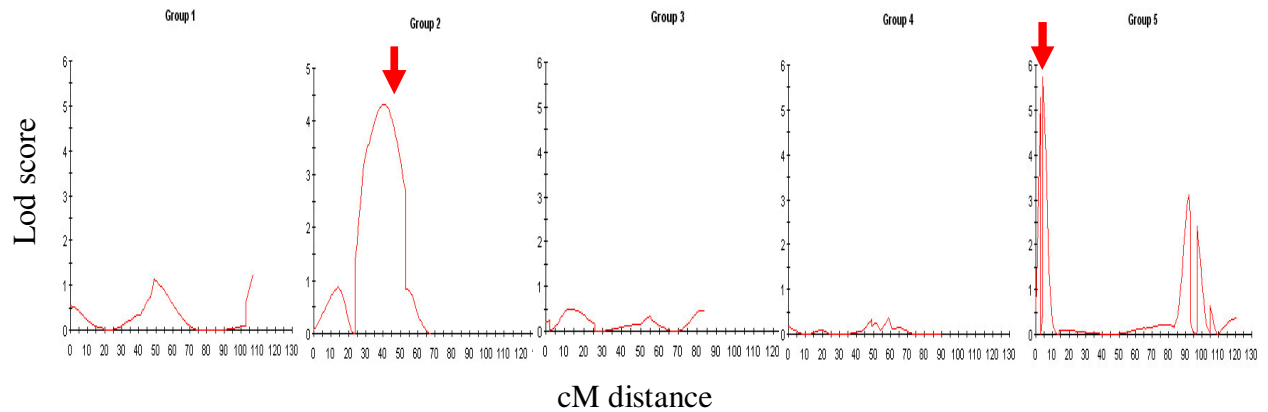


Figure 4.10 QTL linkage group analysis associated with RIL population from Wt-5 x Ct-1 cross. MQM analysis using ‘MapQTL’ software displays peaks at chromosomal locations where trait variance is explained. Significant peak on groups 2 and 5 highlighted by arrowheads.

Chromosome	LOD score	LOD 95% sig. threshold	% variance explained	Sig.QTL	Additive effect
Group 1	1.36	1.8	3.9	No	N/A
Group 2	4.43	1.7	16.7	Yes	+42.9
Group 3	0.51	1.6	1.6	No	N/A
Group 4	0.37	1.6	1.0	No	N/A
Group 5	5.74	1.7	1.1	Yes	+136.0

Table 4.5 Wt-5 x Ct-1 Summary of MQM analysis showing LOD scores and confidence interval with % explanation of variance by QTL location. A positive additive effect indicates that the allele increasing disease was derived from Ct-1 and a negative effect would show that the allele increasing disease was derived from Wt-5. Significant QTLs above 95% statistical confidence level.

4.4.5 Col-0 x Ler

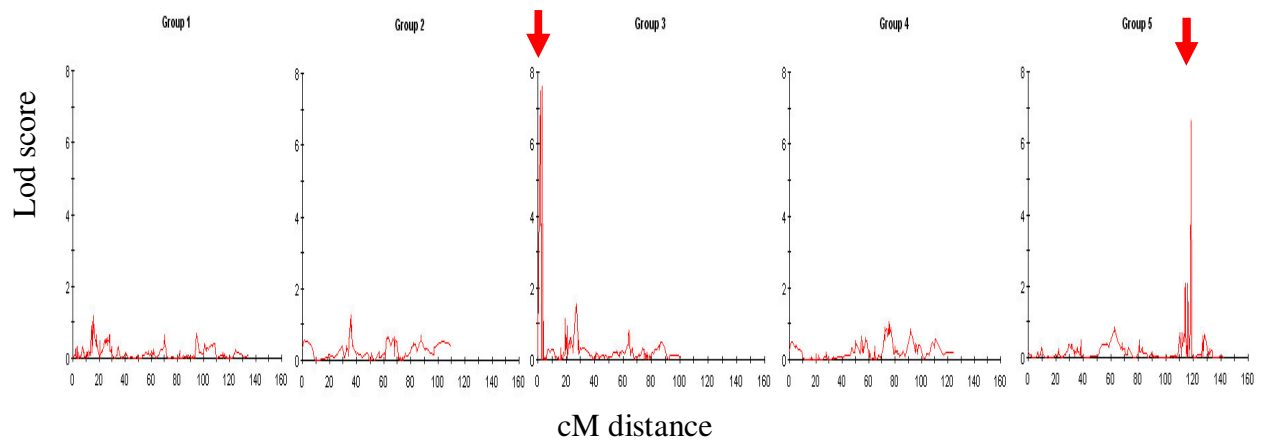


Figure 4.11 QTL linkage group analysis associated with RIL population from Col-0 x Ler cross. MQM analysis using ‘MapQTL’ software displays peaks at chromosomal locations where trait variance is explained. Significant peak on groups 3 and 5 highlighted by arrowheads.

Chromosome	LOD score	LOD 95% sig. threshold	% variance explained	Sig.QTL	Additive effect
Group 1	1.17	2.3	1.8	No	N/A
Group 2	1.25	2.1	2.7	No	N/A
Group 3	7.61	2.1	43.6	Yes	+242.2
Group 4	1.07	2.1	2.0	No	N/A
Group 5	6.73	2.2	16.8	Yes	+27.2

Table 4.6 Col-0 x Ler Summary of MQM analysis showing LOD scores and confidence interval with % explanation of variance by QTL location. A positive additive effect indicates that the allele increasing disease was derived from Col-0 and a negative effect would show that the allele increasing disease was derived from Ler. Significant QTLs above 95% statistical confidence level.

4.4.6 Compiled QTL graphs

The QTL graphs have been compiled to show similarities across the independent populations and analyses.

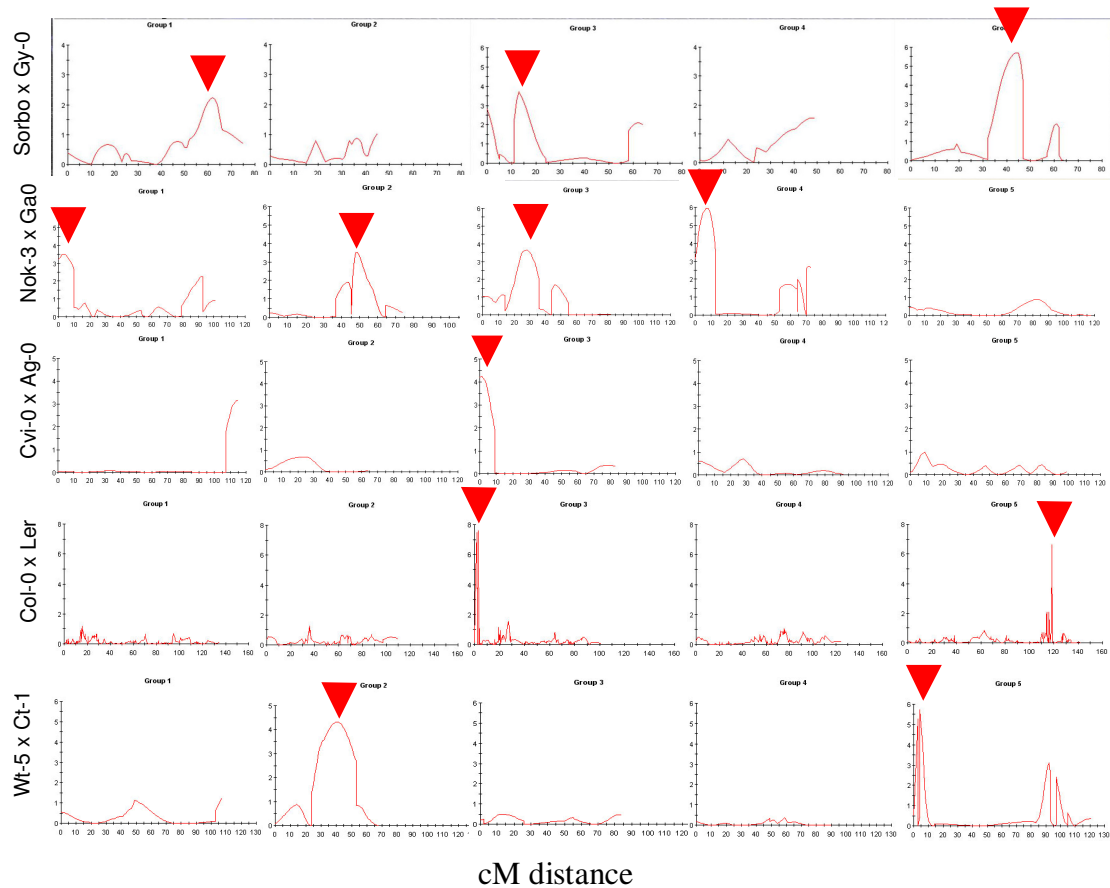


Figure 4.12 QTL linkage group analysis associated with all 5 RIL populations. MQM analysis using ‘MapQTL’ software displays peaks at chromosomal location where trait variance is explained. QTL graphs have been compiled to show similarities across the independent populations. Red arrows represent statistically significant peaks ($P = < 0.05$).

4.5 Selected components define QTL regions

QTL analysis is reliant on marker density, genotypic information and quality /quantity of trait data and is used to generate statistical peak heights above an area of the chromosome which is likely to contribute to the control of the trait. The generated peak is a statistical likelihood of gene location and not an actual location as it lacks the resolution to isolate a specific gene. QTLs and the candidate genes within those areas can be narrowed further by a number of approaches such as the introduction of

more markers and genotype information into the region of interest. The recent improvements in marker density for some RIL populations, in addition to a full sequenced genome for Col-0, allows a different approach which has the potential to narrow candidates down at a faster rate.

All of the genes within the marked QTL regions were included in the initial candidate gene lists. To reduce the possibility of discarding candidate genes, markers with known physical positions were selected to cover the full width of the peak. The initial candidate gene lists are unfeasibly large to investigate and therefore must be refined. Initially, refinement was based on annotation information of reported gene function, protein structure or involvement in disease interaction. A list, based on gene annotation information, of all genes reportedly involved in disease resistance or perception was compiled. The list includes genes obtained from the hierarchical organisation and categorisation of reported gene function from the microarray, 'MapMan' (section 2.4.1). From this software, 421 genes reported to be involved in biotic stress or respiratory burst were obtained. Combined with this list, was a list of genes reportedly involved in PAMP recognition from transcriptional analysis when exposed to bacterial flagellin peptide flg22 (Zipfel et al., 2004). An additional 269 genes were added to the disease resistance gene list from this transcriptional analysis, including families of genes such as the LRR-RLKs (Leucine Rich Repeat-Receptor-Like Kinase) that act as PRRs (Pathogen Recognition Receptors).

Using a multiple interrogation query in Microsoft Access, it was possible to cross-reference this disease-related gene list with QTL derived candidate gene lists. Genes not reportedly involved in biotic stress/ respiratory burst or flg22 stimulated PAMP recognition were eliminated to generate lists with more manageable numbers of genes. For example, Table 4.7 shows the Sorbo x Gy-0 cross has a QTL on chromosome 1 spanning 2227 genes, however after the refinement this is reduced by over 95% to 64 gene candidates that had already been linked to disease resistance, and also located within markers flanking the QTL region detected using the Lux bacteria growth data.

At multiple locations on chromosomes 2 and 3, significant QTL peaks generated from independent analyses have been identified which overlap in their chromosomal area between 5 independent populations. Cross referencing these areas further refined the candidate genes within these marker regions (Fig. 4.12).

RIL population	Chromosome	QTL region (bp)	Closest gene spanning QTL region	Markers spanning QTL region	No. genes within region	No. genes after refinement
Sorbo x Gy-0	1	15926702-23696200	AT1G42470-AT1G63840	F14052-t7p18	2227	64
Sorbo x Gy-0	3	7573001-9796400	AT3G21490-AT3G26650	Mil23i-athgap	643	7
Sorbo x Gy-0	5	7800000-15022000	AT5G23180-AT5G37790	Jconnc-aths01	1650	11
Nok-3 x Ga0	1	8642-3224500	AT1G01020-AT1G09910	Nga590-F16j7	1134	9
Nok-3 x Ga0	2	9934438-16298918	AT2G23340-AT2G39030	Pls888-nga168	1986	52
Nok-3 x Ga0	3	4608284-9697400	AT3G13960-AT3G26483	Nga162-athgap1	1573	18
Nok-3 x Ga0	4	89505-5628809	AT4G00220-AT4G08830	Jv30131-nga888	1354	20
Cvi-0 x Ag-0	3	4608284-7672400	AT3G13960-AT3G21760	Nga163-msd2129380	994	12
Wt-5 x Ct-1	2	8228713-19548527	AT2G18969-AT2G47680	Pls1-t32f646516	3830	109
Wt-5 x Ct-1	5	55123-1572240	AT5G01160-AT5G05310	Mhf15ind52-nga151	1111	8
Col-0 x Ler	3	46403-333029	AT1G01080-AT1G01960	SGCSNP105-SGCSNP379	130	1
Col-0 x Ler	5	24911698-25566425	AT5G62000-AT5G63890	SGCSNP269-SGCSNP29	232	3

Table 4.7 For each of the RIL populations the table describes the base pair locations of significant QTL peaks, the closest gene to the beginning and end of that region as well as the markers defining the region. The total number of candidate genes within the region is listed in addition to the number of refined candidate genes.

4.6 Categorisation of refined candidate genes within the QTL regions

Refined lists of candidate genes found within QTL regions highlighted by all of the different RIL populations were categorised based on annotation information available on TAIR databases (section 2.5.1). Genes were categorised to establish if any single group, family or subcategory of genes or genetic characteristics were prevalent. Groups of candidate genes were assessed separately and as a collective as in Fig. 4.13, which contains genes from all of the QTL areas highlighted in independent populations. Groups of genes with similar characteristics were assessed for the frequency in which they occurred within the QTLs. For the genes which occurred

frequently, a chi squared statistical test determined if they were occurring at significantly higher rates than their background distribution as described in section 4.11. Of the several gene families that occurred frequently within the QTL regions, only the Lysine motif receptor-like kinase gene family was shown to occur at high frequencies in the QTL regions across populations and at significantly high rates compared to background distribution (Table 4.8). The LysM family proteins, predicted to be involved in protein phosphorylation, are located on the plasma membrane and are significantly up-regulated in the presence of a refined protein extract, containing predominantly the flg22 flagellin peptide. These are properties typical in currently known basal resistance gene products (e.g. FLS2, EFR) however, further evidence is required to implicate direct involvement in a basal resistance pathway.

Figure 4.13 Categorisation of refined candidate genes found within QTL regions based on annotation property information across all the RIL populations.

4.7 Correlation between QTL peaks and LysM gene location

Members of the LysM gene family are suggested to be involved in the basal resistance trait from the cross-comparison between QTL peaks and LysM gene location. Figure 4.14 below shows that 9 out of 12 QTL peaks have one or more LysM genes associated with them.

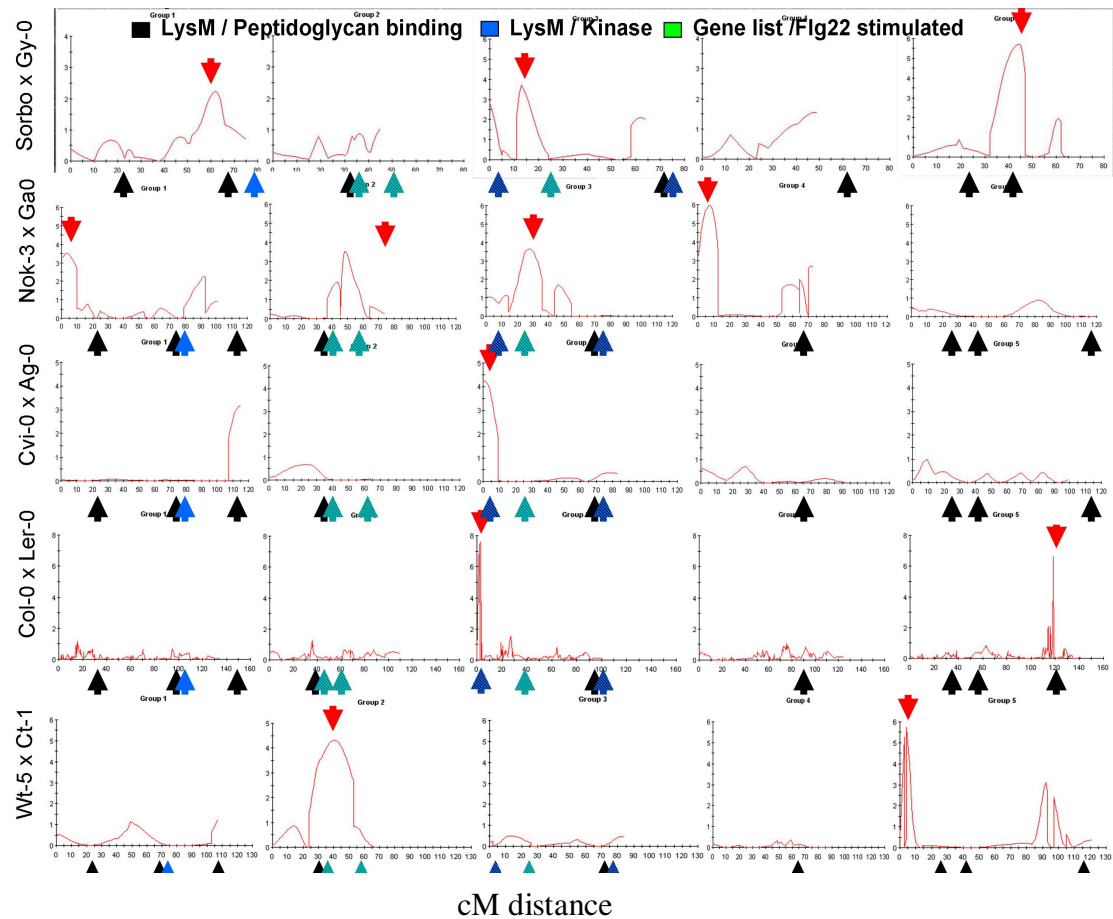


Figure 4.14 Comparison between the QTL peaks and LysM gene locations for 5 independent RIL populations. Red arrows represent statistically significant peaks. Other arrows below the X-axis represent LysM genes with differing colours to categorise their properties.

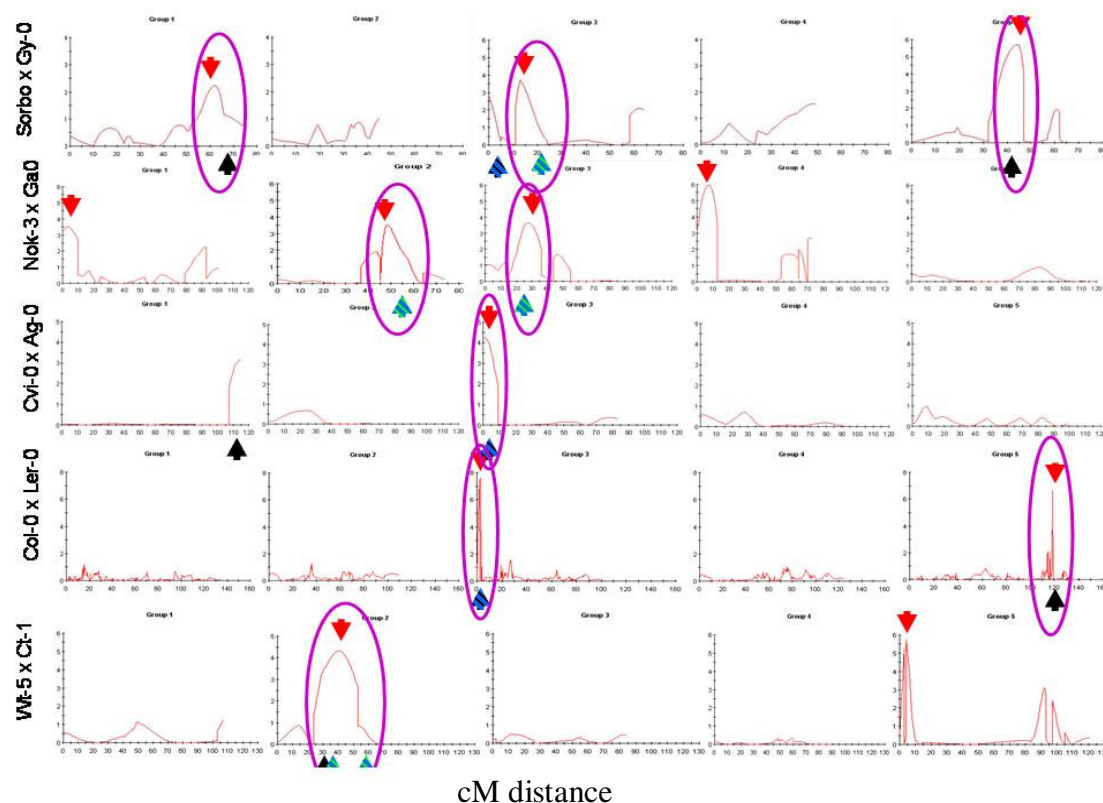


Figure 4.15 A simplified comparison between the QTL peaks and LysM gene locations for 5 independent RIL populations. Of the 9 out of 12 significant QTLs have LysM genes in close proximity to them. Red arrows represent statistically significant peaks. Other arrows represent LysM genes with differing colours to categorise their properties.

4.8 LysM genes distribution statistics

Statistical analysis was used to determine whether the concentration of a particular gene type/ family is significantly over-represented in the QTL intervals. Chi squared matrices compile information including the total gene number, number of genes in the group type of interest, total number of genes found within the QTL regions and number of genes of interest within the QTL regions (Table 4.8). From this information and using Chi squared matrices constructed in the R-2.6 statistical programming language, it is possible to compare the effective hit ratios against the background distribution and prove that the gene of interest is occurring at a significantly higher frequency within the QTL region. The results show that for Col-0 x Ler, LysM genes occur at statistically higher rates in the identified QTL regions compared to the background distribution. The same is true for the majority of the QTL regions associated with the Nok-3 x Ga0 RIL population. The remaining QTL regions do not show significantly higher LysM gene occurrence within QTL regions

compared to background, however, it should be noted that these QTL peaks were less tightly defined to a centimorgan region.

Populat ion	Chromo -some	Chi Matix LysM	Chi Matri x gene No.	Sig.	Populat ion	Chromo -some	Chi Matix LysM	Chi Matrix gene No.	Sig.
Sorbo x Gy-0	Total	3	4520	0.18	Cvi-0 x Ag-0	Total	1	994	0.42
Sorbo x Gy-0	Total	12	20978	0.18	Cvi-0 x Ag-0	Total	14	24504	0.42
Sorbo x Gy-0	1	1	2227	0.25	Cvi-0 x Ag-0	3	1	994	0.24
Sorbo x Gy-0	1	3	4621	0.25	Cvi-0 x Ag-0	3	3	4226	0.24
Sorbo x Gy-0	3	1	643	0.56	Col-0 x Ler	Total	2	362	0.01***
Sorbo x Gy-0	3	3	4577	0.56	Col-0 x Ler	Total	13	25136	0.01***
Sorbo x Gy-0	5	1	1650	0.16	Col-0 x Ler	3	1	130	0.02***
Sorbo x Gy-0	5	2	4224	0.16	Col-0 x Ler	3	3	5090	0.02***
Nok-3 x Ga0	Total	4	6047	0.06*	Col-0 x Ler	5	1	232	0.03***
Nok-3 x Ga0	Total	11	19451	0.06*	Col-0 x Ler	5	2	5642	0.03***
Nok-3 x Ga0	1	1	1134	0.10*	Wt-5 x Ct-1	Total	3	3830	0.42
Nok-3 x Ga0	1	2	4913	0.10*	Wt-5 x Ct-1	Total	12	21668	0.42
Nok-3 x Ga0	2	1	1986	0.42	Wt-5 x Ct-1	2	3	3830	0.31
Nok-3 x Ga0	2	2	4061	0.42	Wt-5 x Ct-1	2	0	207	0.31
Nok-3 x Ga0	3	1	1573	0.18					
Nok-3 x Ga0	3	2	3647	0.18					
Nok-3 x Ga0	4	1	1354	0.07*					
Nok-3 x Ga0	4	2	4693	0.07*					

Table 4.8 Information used in the Chi-squared matrices to assess LysM gene distribution. Figures marked with an asterisk have a significant difference of <0.05 indicating LysM distribution occurring at a significantly higher frequency within QTL region compared to the background. Significance threshold value – *0.1 **0.05 ***0.01 ****0.005 *****0.001.

4.9 Properties of the LysM gene family

Cross comparison of candidate gene lists from QTL peaks implicated a *LysM* gene family member, which occurs frequently in QTL highlighted regions across populations (Fig.4.15). FLS2 is a known PAMP receptor which is associated with the detection of bacterial flg22 and has very similar properties to LysM. Property similarities include membrane associated Serine/ Threonine and Tyrosine protein kinase domains for the phosphorylation of the target protein substrate. These properties are conserved across the LysM gene family, with the notable property difference between LysM gene products and FLS2 being the absence of LRR domains. Therefore, LysM share many attributes of known PAMP receptors and appear mechanistically capable of acting as PAMP receptors with domains that interact with chitin and peptidoglycan (Zhang et al., 2007; Buist et al., 2008; Gimenez-Ibanez et al., 2009). Apart from a small number of specific examples, the mechanisms of action for PAMP receptors in basal resistance are not known. The product of the LysM RLK gene *CERK1* has been shown to bind with fungal chitin and instigate a basal defensive signal cascade, however, there is also evidence that CERK1 may also bind to peptidoglycan, which is a critical component in bacterial cell walls (Zhang et al., 2007; Buist et al., 2008; Gimenez-Ibanez et al., 2009). Property similarities are only an indication of possible mechanistic overlap however, direct proof of LysM genes involvement in bacterial perception for basal resistance is needed.

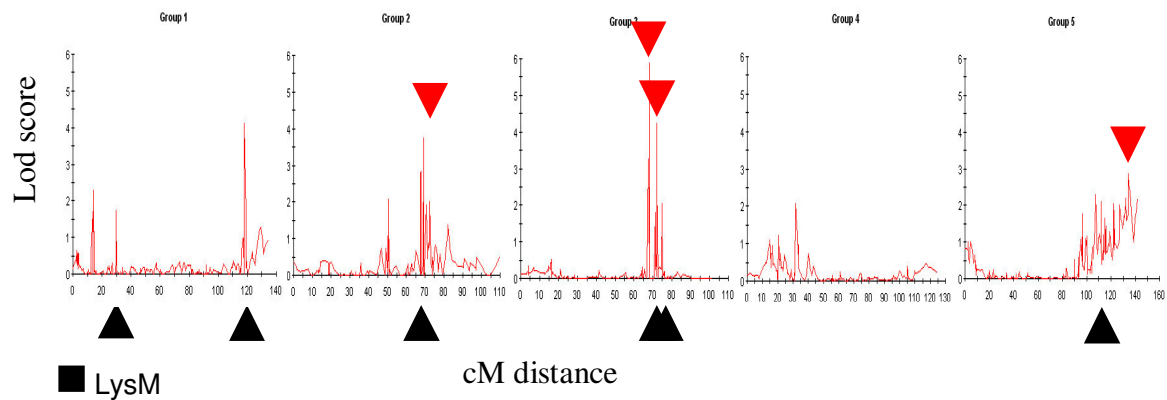
4.10 Reanalysis of QTLs following *Psm* infiltration

The QTLs and/ or the LysM genes implicated to be involved in basal resistance have been derived from a specific interaction between *Pst* and *Arabidopsis*. Retesting 2 RIL populations with *Pseudomonas syringae* pv. *maculicola* (*Psm*) will determine whether the same QTLs determine basal resistance to different *Pseudomonas* pathovars. The growth conditions, experimental design, scoring etc. were all kept the same as those used in the *Arabidopsis* / *Pst* trials, with the exceptions that the bacterial growth was measured 6 hours earlier than in *Pst* trials to compensate for the increased tissue damage of *Psm* ES4326 strain over *Pst* DC3000 at the typically measured time point (48hrs) (Davis et al., 1991). Figures 4.16 and 4.17 show *Psm*

QTL graphs and the relative positions of *LysM* genes for comparison with *Pst* QTL graphs. In this QTL analysis, only the trait source data is different (*Pst* and *Psm*). The genotype information and marker density are the same therefore QTL resolutions are comparable.

4.11.1 Col-0 x Ler *Psm*

a) Col-0 x Ler *Psm*



b) Col-0 x Ler *Pst*

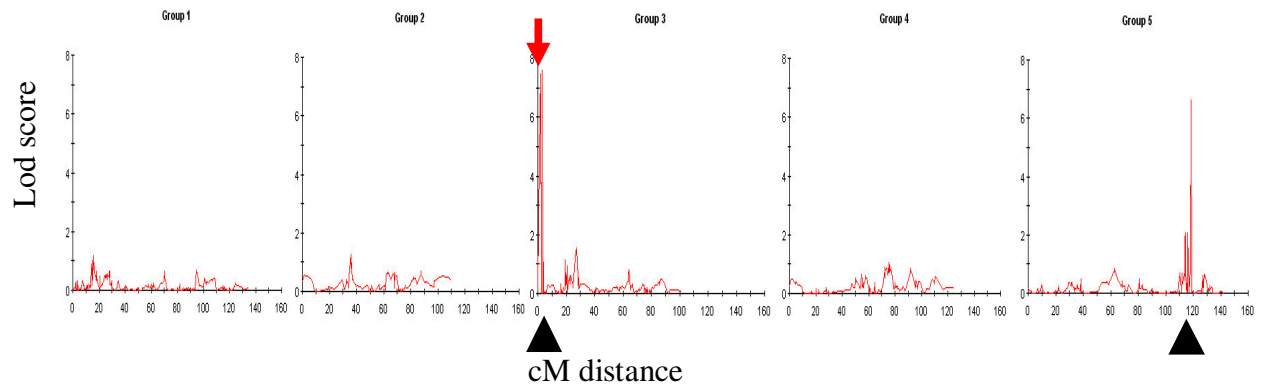


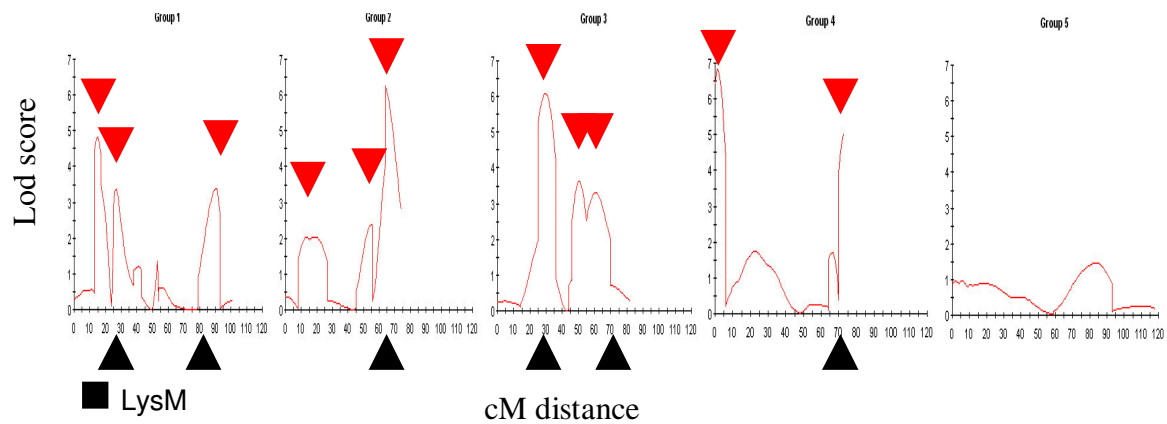
Figure 4.16 MQM analysis using ‘MapQTL’ software displays peaks at chromosomal location where trait variance is explained. (a) QTL linkage group analysis associated with RIL population from Col-0 x Ler infiltrated with OD₆₀₀ 0.002 *Psm*. (b) Equivalent data for Col-0 x Ler infiltrated with OD₆₀₀ 0.002 *Pst*. For both (a) and (b) red arrows represent statistically significant peaks and black arrow show reported *LysM* gene locations ($P < 0.05$).

Chromosome	LOD score	LOD 95% sig. threshold	% variance explained	Sig. QTL	Additive effect
Group 1	4.11	2.3	15.1	Yes	+54.3
Group 2	3.75	2.2	15.4	Yes	+35.1
Group 3	5.89	2.2	25.9	Yes	+53.8
Group 3	4.1	2.2	15.0	Yes	+50.7
Group 4	2.06	2.1	9.7	X	N/A
Group 5	2.86	2.2	13.2	Yes	-12.6

Table 4.9 Col-0 x Ler Summary of MQM analysis of Col-0 x Ler showing LOD scores and confidence interval with % explanation of variance by QTL location. A positive additive effect indicates that the allele increasing disease was derived from Col-0 and a negative effect would show that the allele increasing disease was derived from Ler. Significant QTLs above 95% statistical confidence level were detected on groups 1, 2, 3 and 5.

4.11.2 Nok-3 x Ga-0 *Psm*

a) Nok-3 x Ga-0 *Psm*



b) Nok-3 x Ga-0 *Pst*

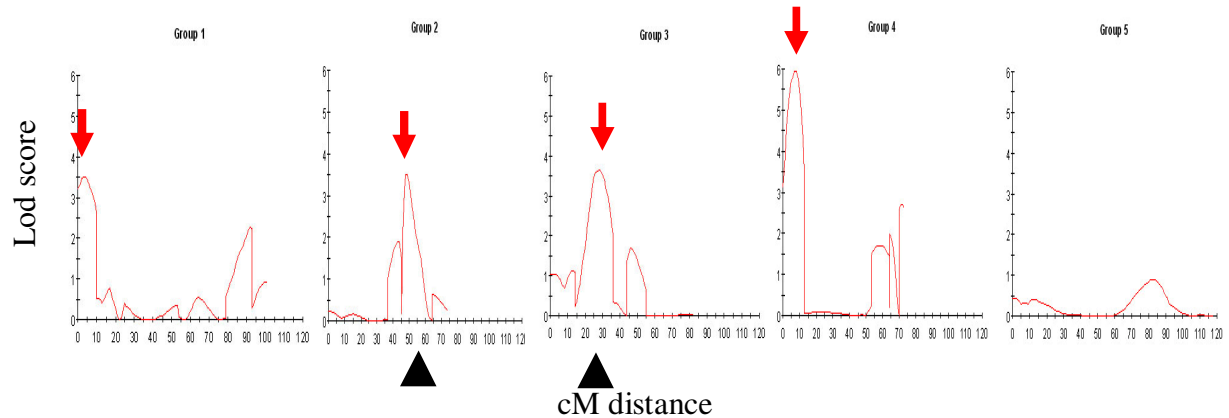


Figure 4.17 QTL linkage group analysis associated with RIL population from Nok-3 x Ga-0 infiltrated with OD₆₀₀ 0.002 *Psm*. MQM analysis using ‘MapQTL’ software displays peaks at chromosomal locations where trait variance is explained. Significant peaks on groups 1, 2, 3 and 4. Red arrows represent statistically significant peaks and black arrows show reported LysM gene locations. Graph (b) reproduces the equivalent data for Nok-3 x Ga-0 infiltrated with OD₆₀₀ 0.002 *Pst* (Fig. 4.8) for direct comparison.

Chromosome	Location (cM)	LOD score	Lod 95% sig. threshold	% variance explained	Sig. QTL	Additive Effect
1	15	4.82	2.7	9.4	Yes	-34.5
1	27	3.38	2.7	12.1	Yes	+101.6
1	90	3.39	2.7	6.9	Yes	-29.8
2	64	6.26	1.8	13.6	Yes	-25.8
2	55	2.37	1.8	5.4	Yes	+37.8
2	18	2.04	1.8	3.9	Yes	-58.2
3	30	6.08	1.8	15.3	Yes	+73.4
3	51	3.62	1.8	8.7	Yes	+51.5
3	61	3.31	1.8	8.4	Yes	+48.4
4	1	6.84	1.8	15.7	Yes	+46.8
4	73	5.04	1.8	10.5	Yes	+48.6
5	83	1.46	1.7	3.3	No	N/A

Table 4.10 Nok-3 x Ga-0 Summary of MQM analysis showing LOD scores and confidence interval with % explanation of variance by QTL location. A positive additive effect indicates that the allele increasing disease was derived from Nok-3 and a negative effect shows that the allele increasing disease was derived from Ga0. Significant QTLs above 95% statistical confidence level detected on groups 1, 2, 3 and 4.

4.12 *AvrPtoB*

The *AvrPtoB* is a type III effector protein expressed by the wild-type *Pst* DC3000 bacterial strain which exhibits E3 Ub ligase activity (Abramovitch et al., 2006b; Janjusevic et al., 2006). This bacterial effector enhances pathogenicity as it acts to ubiquitinate the Ser/Thr kinase receptor domain of host R-gene-encoded Pto proteins and, by the same mechanism, suppresses the kinase function of FLS2 in Col-0 Arabidopsis by proteasomal degradation (Abramovitch et al., 2003; Craig et al., 2009). Based on their annotated sequence information, six of the LysM genes contain a Ser/Thr receptor-like kinase domain similar to that found in the known PAMP receptor FLS2 (see section 4.8). Preliminary data using tomato orthologues show that *AvrPtoB* secreted by *Pst* may suppress PAMP triggered immunity and that CERK1 is the possible host target. The *AvrPtoB* effector appears to have broad ranging Ser/Thr kinase domain suppressive function (Lin and Martin, 2005). The experiment described below aimed to assess the extent of any potential *AvrPtoB* suppression effect shown in *R*-genes and FLS2 PAMP receptors for a variety of RIL ecotypes and to determine if a correlation could be shown between QTLs, LysM gene location and

possible suppression of kinase function, resulting in a difference in bacterial proliferation. Infiltration of wild-type *Pst* already containing *AvrPtoB* or the deletion mutant *Pst* Δ *AvrPtoB* into a selection of ecotypes was measured using a dilution plating assay as the mutant strain has not been Lux tagged (Lin and Martin, 2005). The results (Fig.4.18) show a non-significant trend for slightly increased bacterial growth for the majority of genotypes when infiltrated with wild-type *Pst* DC3000, compared with the deletion mutant of *Pst* Δ *AvrPtoB*. With the exception of Nok-3 T-test which shows a significantly higher ($P = <0.05$) bacterial count in the *Pst* containing *AvrPtoB* compared to the deletion mutant, the other ecotypes are not significantly different. Some genotypes, such as Ct-1, display very little difference between infiltration strains and others show, such as Ga-0, the reverse to the general trend significantly ($P = <0.05$). F-Tests compared *Pst* to the *AvrPtoB* mutant within a ecotype and showed no significant differences in the variance for any of the ecotypes ($F_{pr} >0.05$). The results do not show a clear pattern between QTLs, LysM gene location and suppression of kinase function however, the observations suggest *AvrPtoB* may have a role in suppression of basal resistance for some genotypes, but that conflicting influence factors may also be involved.

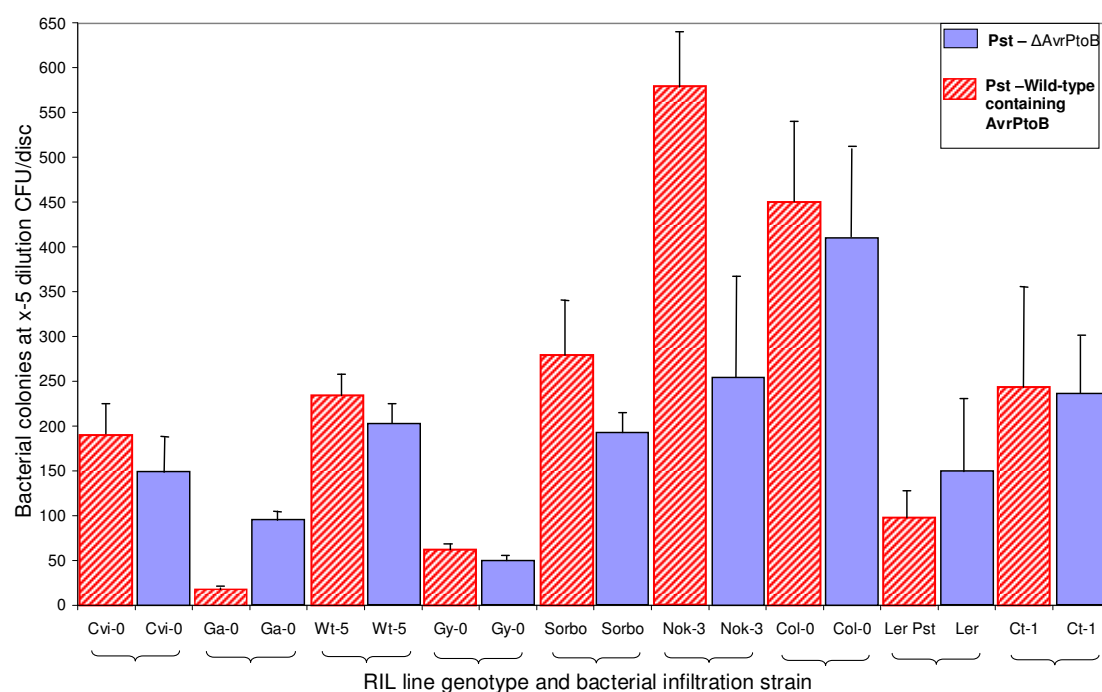


Figure 4.18 Comparison of growth of wild-type *Pst* DC3000 (containing *AvrPtoB*) and the *Pst* deletion mutant of *AvrPtoB* (*Pst*- Δ *AvrPtoB*) in 9 parental RIL genotypes measured using dilution plating and colony counting. The error bars represent standard deviation (SD).

4.13 Discussion

4.13.1 QTL Analysis

Five distinct RIL populations were chosen from the parental line trial which differed in their basal resistance. Bacterial growth of *Pst* was measured in 94 RILs derived from original crosses of: Sorbo x Gy-0, Nok-3 x Ga0, Cvi-0 x Ag-0, Col-0 x Ler and Wt-5 x Ct-1 (Fig.4.2-4.6). Lux trait data trials showed statistically significant differences between RILs and no significant variation between replicate plants of a RIL line. Replicate trays were compared and showed no significant differences between them, as anticipated for the newly refined assay procedures. The trait data from these independent RIL lines all follow a very similar pattern with a continuous range of variation suggesting a multigenic control over basal resistance. The variation range within a RIL population extended in excess of 50-fold for lines such as Col-0 x Ler and to over 200-fold for the Nok-3 x Ga0 line, with the other populations falling amid these points. These ranges are primarily due to the differential efficacy of basal mechanism to control *Pst* proliferation as there is no *R* gene action (Fan et al., 2008). Reductions in the standard deviation around the mean for Lux trait data from improvements to the procedure increased confidence that the assay could accurately differentiate between RIL lines. Therefore, QTL analysis could more accurately define the chromosomal location influencing the trait.

QTL analysis aims to define the number and location of the component differences in basal mechanisms between RIL lines and ecotypes which act to alter the trait. Significant peaks indicate areas on the chromosome which contain gene(s) which alter the trait and 12 were found at multiple locations from across the RIL populations. Nok-3 x Ga0 population showed the largest range of trait scores and had the highest number of significant peaks, 4 on 4 different chromosomes. This is in contrast with Col-0 x Ler which had only 2 QTL peaks on 2 chromosomes and the lowest variation in trait. Other RIL populations' trait variation and numbers of significant QTLs fell roughly between these extremes. The trend then suggests that an increase in the number of significant QTLs correlates with an increase in the range of a RIL populations' trait variation.

Individually, the graphs indicate the chromosomal regions associated with basal resistance. The Sorbo x Gy-0 RIL population revealed significant QTLs at loci

around 60cM, 20cM and 40cM on chromosomes 1, 3 and 5 respectively, which explain around 80% of the variation. Subsequent analysis on the Ga-0 x Nok-3 RILs revealed significant QTL peaks on chromosomes 1, 2, 3 and 4. A significant QTL was located on chromosome 3 in the Cvi-0 x Ag-0 RILs and two tightly defined QTL peaks were found on chromosomes 3 and 5 in the Col-0 x Ler RILs. In addition, significant QTL peaks were found on chromosomes 2 and 5 for the Wt-5 x Ct-1 RIL population. The additive effect shows, as expected, that the variation at the QTL locations is mainly derived from the most susceptible parental line for each of the RIL populations. Compilation of the graphs from independent RIL populations showed significant QTL peaks overlapping in similar genetic locations which implicate these chromosomal regions in the underlying control of basal resistance. There are three overlapping QTL regions between Sorbo x Ga-0 and Nok-3 x Ga0, Cvi-0 x Ag-0 and Col-0 x Ler, Nok-3 x Ga0 and Wt-5 x Ct-1 as shown in Figure 4.15. This suggests that genes contained within the QTL regions are likely to be conserved across the *Arabidopsis* population. QTL analysis normally does not have the resolution to isolate specific genes however, it is a useful early stage of analysis that can be used to generate an initial candidate gene list based on the chromosomal areas implicated.

4.13.2 Candidate gene refinement

In studies using a similar QTL approach for a single highly significant QTL, it has been common to introduce further markers and genotypic information into this region to improve QTL resolution and refine candidate gene numbers (Clarke et al., 1995). Typically, this approach can narrow down the identified chromosomal area however, it is time intensive and can usually only reduce candidates to, at best, approximately 50 genes before maximum QTL resolution is reached (Clarke et al., 1995). There has recently been a dramatic increase in the available marker and genotype information available for the Col-0 x Ler lines as both have now been fully sequenced. As a result, marker density in these lines is high (~1300 across the genome) and corresponding QTL resolution is approaching its peak and so additionally added markers would only produce modest refinement of candidate genes. The other RIL lines have lower marker densities however the constraints of this project made the time-intensive introduction of more markers into the regions an inefficient use of time

with only marginal gains expected in QTL resolution. A more direct approach was taken of selecting markers either side of each of the significant QTL peaks and listing the genes contained within those areas as initial potential candidates. Using annotation information and classifications based on the physiological or pathway association from the 'MapMan' software, the initial candidate gene lists were refined and reduced by approximately 95%. For example, the 2227 genes within the markers of Sorbo x Ga-0 chromosome 1 QTL were reduced to 64 candidates and the 643 genes in the QTL on chromosome 3 for the same cross were reduced to 7 candidates (Table 4.7). The initial lists were too large to be effectively screened therefore candidate genes were refined however the risk of discarding genes unnecessarily needed to be minimised. Candidate genes that had no published association to plant disease or biotic stress were discarded and the remaining candidates were categorised based on properties (Fig.4.13). This would establish if any particular gene family or characteristic is prevalent within QTL regions based on annotation information (e.g. Lectin-RLKs, Protease inhibitors or Nucleotide Binding Site). As long as at least one member of a family has been implicated then it was considered a potential candidate. It should be acknowledged there remains a small risk of discarding genes unnecessarily which have no previously identified link to plant disease and other gene groups with similar ontologies that do not occur in significant high rates in the identified QTL regions. However, the combination of these two refinements minimises this small remaining possibility.

A cross-comparison of refined candidate lists from each of the independent populations found multiple regions implicating similar genes in more than one population. From this approach, two LysM genes found within the region implicated by QTL analysis were initially noted as prominent candidates. Members of the LysM gene family were mapped relative to the locations of the QTLs and a striking correlation was observed. Nine out of 12 QTLs line up directly with LysM genes and 9 out of 15 LysM genes fall within QTL regions. This correlation was sufficiently compelling to form the hypothesis that the LysM gene family had an active role in basal resistance functional pathways. To test this correlation, multiple chi squared matrices were constructed to compare the effective hit ratios in the QTL regions with the chromosomal background distribution. Other genes / families were tested such as Serine protease inhibitors, CC-NBS-LRRs and TIR-NBS-LRRs but only LysM genes showed some significant results. Two LysM genes implicated in separate regions of

the Col-0 x Ler population showed >95% statistically significant differences (Table 4.8). LysM genes associated with the Nok-3 x Ga0 cross as a total shows a P-value of 0.1 association with QTL peaks, which approaches the standard 95% significance threshold. This evidence suggests that the LysM gene family does have a significant association with QTLs implicating areas controlling basal defence and therefore could be active components. The marker density in the Col-0 x Ler population is substantially higher than other populations therefore, the markers could refine the peaks to a narrower chromosomal area so background variation was reduced. If the other populations had such high marker densities, it is likely that they would also show significant results as many were already approaching significance (Table 4.8). During marker selection either side of the QTL peaks, a strategy of wide marker positioning to avoid discarding potentially important genes works against the statistical tests because it includes more background into the QTL region. Despite the limited success rate for the statistics, there remains sufficient evidence for the LysM gene family involvement to warrant further examination.

4.13.3 Specificity of the basal resistance conferred by the QTLs

The initial survey of RIL populations derived from independent genotypes was conducted using *Pst* however, this may limit findings if similar trends cannot be observed in other bacterial strains and species. Retesting 2 RIL populations with *Psm* aimed to determine whether the QTLs, and potentially the gene function, are conserved across bacterial species as part of a broader perception of pathogens or whether the QTLs are race specific. Figures 4.16 and 4.17 show the relative positions of LysM genes for comparison between *Pst* and *Psm* QTL graphs. In the Nok-3 x Ga-0 RIL populations both *Pst* and *Psm* infiltration and subsequent QTL analysis both show significant peaks at ~10cM on chrom1, ~60cM on chrom2, ~30cM on chrom3 and ~5cM on chrom4. The peak profiles between *Pst* and *Psm* are very similar however the *Psm* infiltrated QTL analysis shows additional areas of significance at around 90cM on chrom1, ~20cM on chrom2, ~60cM on chrom3 and ~70cM on chrom4 where the *Pst* peaks only approach significance at those locations. The increased virulence of *Psm* ES4326 strain over *Pst* DC3000 at the typically measured time point (48hrs) creates a greater contrast between susceptible and resistant RIL line

bacterial measurements. This aids the QTL analysis software to distinguish additional areas of significance in the *Psm* infiltrated compared to *Pst* (Davis et al., 1991). Comparison of Nok-3 x Ga-0 RIL populations QTL graphs show substantial similarity between the QTL peak profiles. This suggests the same chromosomal regions are involved in basal perception of both *Pseudomonas* bacterial pathogens and a broadly similar basal target/ receptor interaction between *Pst* and *Psm*. In contrast, cross-comparison between *Pst* and *Psm* infiltrated lines for the Col-0 x Ler populations shows a different pattern where there is very little similarity between the two sets of QTL peak profiles. This result may imply different perception mechanisms conferring resistance to *Pst* and *Psm*. However, there is still a close association with LysM genes for both of the *Pst* and *Psm* interactions which still implicates LysM genes as being involved in perception of a broad bacterial pathogen range. The PAMP signal from *Pst* and *Psm* is likely to be the same as molecular components and are typically highly conserved within a class of microbes, therefore, the current hypothesis implies that LysM gene receptor combinations confer recognition specificity, however further experimentation would be needed to test this hypothesis (Nicaise et al., 2009).

4.13.4 Role of LysM gene family in basal disease resistance

Comparison of LysM genes with known PAMP receptors such as FLS2, shown to be involved in the perception of bacterial flg22, was conducted (Navarro et al., 2004; Zipfel et al., 2004). Annotation information described several similar properties between LysM genes and the known PAMP receptors for example, they share predicted serine/ threonine kinase domains and are predicted to act as transmembrane receptors. Major functional differences between LysM genes and the known PAMP receptor FLS2 included the absence of Leucine rich repeat domains. LRR domains are typically important in offering a structural link between binding sites so that binding sites are orientated correctly to bind with their substrate and thus it can be critical to gene function (Kobe and Kajava, 2001).

LysM genes found in *Arabidopsis* share strong homology to the *NFR1* gene which acts as a receptor involved in rhizobial Nod factor perception in *Lotus japonicus*. It has been shown that a combination of *NFR*-type genes can confer

specificity to perceive *Rhizobium* and initiate legume nodulation formation (Radutoiu et al., 2003). Strong homology between the genes (LysM family and *NFR1* and *NFR5*) introduces the possibility of similarity between mechanisms in Nod factor and basal perception pathways. These properties are consistent with the hypothesis of LysM involvement in known mechanisms of PAMP and so are worth pursuing.

Recently the LysM RLK gene *CERK1* has been implicated to be involved in the perception of chitin, a major component of the fungal cell wall (Kaku et al., 2006). Mutations in this gene render the host plant completely susceptible to fungal infection and are necessary for the basal signalling and defence activation via the WRKY and MAPK mediated pathway (Miya et al., 2007). A homologues *CERK1* gene found in rice, *CEBiP*, is also involved in binding to chitin and knockdown transformants show increase susceptibility and decrease defence activation when exposed to fungal pathogens (Kaku et al., 2006). The QTL analysis implicating *LysM* genes when exposed to bacterial pathogens suggests that *LysM* genes may not be specific to fungal compound recognition (Gimenez-Ibanez et al., 2009). This hypothesis is supported by evidence showing that LysM domains are able to bind to peptidoglycan, which is a primary component of bacterial cell walls (Buist et al., 2008).

4.13.5 Effect of *AvrPtoB* on bacterial growth

AvrPtoB is a highly conserved bacterial effector which has been previously implicated in plant defence through inhibition of programmed cell death by suppression of intermediates such as Pto and Cf9 (Abramovitch et al., 2003). Preliminary research using the tomato LysM orthologue of the *CERK1* gene, implicated *AvrPtoB* as a potential inhibitor which increased bacterial virulence by suppression of basal defence. This hypothesis was expanded and shown to be applicable to *Arabidopsis* by demonstrating that *AvrPtoB* targets the CERK1 LysM receptor kinase to increase host susceptibility (Abramovitch et al., 2006a; Gimenez-Ibanez et al., 2009). However, the specificity of a CERK1-*AvrPtoB* interaction is complicated by the observation that the PAMP receptor FLS2 is also directed for degradation by *AvrPtoB* (Goehre et al., 2008). *AvrPtoB* associates with FLS2 and contributes to virulence, probably through degradation by the attachment of ubiquitin

molecules to the Ser/Thr kinase domain of FLS2, which may suggest broad ranging suppression of receptor-like kinase components (Goehre et al., 2008).

The experiment aimed to determine any effect *AvrPtoB* contained within the *Pst* DC3000 wild-type has on bacterial growth compared an *AvrPtoB* deletion mutant. *AvrPtoB* reportedly targets CERK1 via suppression of kinase function, so if this interaction is specific, then it is possible that a correlation could be observed between genotypes with QTLs implicating *CERK1* region and bacterial growth alteration (Zipfel and Rathjen, 2008). A specific kinase suppression hypothesis would suggest that RIL lines with significant QTLs involving the *CERK1* gene would show a significant difference when comparing bacterial growth of *Pst* with or without *AvrPtoB*. Consequently, little difference in bacterial growth would be observed in other genotypes which implicated different LysM genes. Bacterial growth data for multiple genotypes previously used to generate QTLs, which implicate the *CERK1* region, were tested (in addition to those which do not) to establish any correlation linking *CERK1* QTLs and suppression pattern.

The plate assays showed a trend towards increased bacterial growth when *AvrPtoB* was present for 7 out of 9 genotypes (typical increase of ~5-20%). For the majority of ecotypes, this was not a significant difference ($P = >0.05$), however Nok-3 did show significantly higher bacterial growth in the *Pst* line with *AvrPtoB*. For some ecotypes, *AvrPtoB* may be a factor involved in the suppression of basal defence as the only difference between the *AvrPtoB* mutant and the wild-type *Pst* was a deletion of 2,007-bp sequence, encompassing most of *AvrPtoB* and a small region downstream (Lin and Martin, 2005). The bacterial growth assays showed no clear correlations between QTLs and *AvrPtoB* suppression pattern, and this may suggest any *AvrPtoB* suppression is not limited to *CERK1* and may have suppressive kinase function which affects multiple LysM genes. Previously reported *AvrPtoB* suppression of FLS2 and the *R*-gene *Pto* would support this hypothesis and suggest broad kinase suppression, which would explain the increased susceptibility trend for the majority of genotypes tested (Lin and Martin, 2005). However this trend is not universally applicable as 2 lines (Ga-0 and Ler) showed the reverse trend and some lines only showed mild differences which suggests confounding factors. *AvrPtoB* is delivered into *Arabidopsis* genotypes through type III secretion, however this system is open to genotypic interference as different plant lines have variation in susceptibility to type III secretory mechanisms, consequently influencing the results (Collmer et al., 2000;

Fouts et al., 2003). *AvrPtoB* in the wild-type *Pst* does significantly affect basal resistance in comparison with the deletion mutant in the Nok-3 ecotype, however other ecotypes showed a ~5-20% which was a not statistically significant increase in susceptibility. Comparison between wild-type DC3000 and the *AvrPtoB* deletion mutant were published for the Col-0 and Ws-4 ecotypes and showed small increases in virulence for the DC3000 with a weak significance ($P = <0.01$) for Col-0 but a stronger ecotype increase difference using Ws-4 ($P = <0.05$) (Abramovitch et al., 2006a; Gimenez-Ibanez et al., 2009). Consistently significant differences of $P = <0.05$ required the double mutant deletion of *AvrPto* and *AvrPtoB* (Janjusevic et al., 2006; Craig et al., 2009). This is probably due to broad kinase suppression and therefore it is likely that at some point in the basal resistance pathway, a kinase gene is involved. Whether kinase suppression targets one or multiple LysM genes and whether they are even involved directly in basal resistance requires further evidence.

Chapter 5: LysM gene family characterisation

5.1 Introduction

The distribution of the *LysM* gene family in relation to the identified QTLs in the RIL populations suggested members of this family play a role in basal resistance. A working hypothesis was formulated from advances in the literature, combined with experimental evidence to predict a possible mode of action or role for members of the *LysM* gene family within basal disease resistance. Transcriptional, sequence, phylogenetic and knockout experiments were utilised, aiming to test this hypothesis.

The LysM gene family consists of 15 membrane-bound genes with similar properties which are spread evenly over the 5 *Arabidopsis* chromosomes, however, a smaller subset of 6 genes are predicted to have RLK function (Fig. 5.2). The LysM domain consists of a $\beta\alpha\alpha\beta$ secondary structure with two α -helices flanked by two-stranded anti-parallel β -sheets (Buist et al., 2008). In the LysM gene family, only the *CERK1* (with RLK function) has currently been implicated in pathogen defence through its role in fungal chitin oligosaccharides perception (Wan et al., 2008). This induces a cascade which results in enhanced expression levels, compared to wild type, of WRKY transcription factors and MAPKs signalling that are established components of *R* gene / basal response (Zhang et al., 2007; Wan et al., 2008). Mutants of the *CERK1*, gene possessing a dominant mis-sense mutation in the LysM domain, show fully compromised non-host resistance to fungal infection, therefore, strongly associating *CERK1* with fungal perception and defence (Lipka et al., 2007). Further evidence of LysM association with fungal perception is shown in an orthologue to *CERK1*, the chitin elicitor-binding protein (CEBiP) from rice. RNAi silencing of this plasma membrane bound glycoprotein with LysM motifs shows reduces chitin binding and increased susceptibility (Miya et al., 2007; Zipfel, 2009).

In contrast to fungal perception association, preliminary data shows that AvrPtoB kinase inhibitor secreted by *Pseudomonas* bacteria may suppress PAMP triggered immunity in tomato orthologues and that host CERK1 is a possible target for defence suppression (Miya et al., 2007; Gimenez-Ibanez et al., 2009). *CERK1* displays similar expression patterns during exposure to both chitin and peptidoglycan. Combined with the evidence that the LysM domain may actively bind to peptidoglycan, a major component of bacterial cell walls, this suggests a broader

perception mechanism (Buist et al., 2008). A LysM receptor based mechanism may target a conserved structural component found in both fungal and bacterial cell walls such as a monosaccharide or a small polysaccharide chain (Buist et al., 2008).

NFR1 and *NFR5* from *Lotus japonicus* both contain LysM-type serine/threonine receptors and work in conjunction to detect extracellular lipo-chito-oligosaccharides bacterial Nod factor signals in the roots (Fig.5.1) (Madsen et al., 2003). Knockout mutants of *NFR1* and *NFR5* are both required for successful signal transduction which indicates a likely receptor complex or alternatively, a close association where the two independent signal-transduction pathways converge after Nod factor perception (Lohmann et al., 2010). Transfer of the *L. japonicus* *NFR1* and *NFR5* into *Medicago truncatula* extended host range to include modified strains of *Mesorhizobium loti* and *Rhizobium leguminosarum* which specifically target *L. japonicus* (Lohmann et al., 2010). This evidence demonstrates the combination of NFR receptors confers host specificity and gene modification experiments show the second LysM motif of *NFR5* is a major determinant in Nod factor recognition (Lohmann et al., 2010). *NFR1* shows strong sequence homology to the *Arabidopsis CERK1* gene (as described in section 5.3.2) which suggests that functional similarity and a possible evolutionary relationship may exist between Nod factor and basal resistance mechanisms (Fig.5.1) (Wan et al., 2008). The evolution and regulatory relationships of the LysM receptor gene family have been mapped for *Lotus japonicus* using a combination of phylogenetic analysis and gene structure comparison (Lohmann et al., 2010). This analysis revealed a common phylogenetic source and the family is likely to have expanded by a combination of tandem and segmental duplication events as a result of positive selection pressure to optimise ligand interaction with the predicted Nod-factor binding domain (Lohmann et al., 2010). Microarray expression analysis showed some *Lotus japonicus* LysM genes up-regulated during both the symbiotic association with *Mesorhizobium loti* and in response to chitin treatment (Lohmann et al., 2010). The diversification of the LysM family and their interaction to confer specificity in legumes is hypothesised to be a result of the pressure imposed by an increasing diverse array of chitin-derived molecules, produced by an extended range of interacting organisms with differing life-cycle objectives. The pressure on legumes to accurately differentiate between potentially symbiotic and pathogenic bacteria forced the evolution and specificity of the current *NFR1*, *NFR5* and other *Lotus* LysM genes (Lohmann et al., 2010). LysM

gene encoded receptors may operate in a similar model to *NFR1* and *NFR5*. Therefore be working in conjunction to target compounds which have structurally similar domains to NOD factor such as lipopolysaccharides, lipo-chito-oligosaccharide, peptidoglycan, β -glucans and chitin oligosaccharides, and which are found in both fungi and bacteria (Buist et al., 2008). The hypothesised mechanistic model predicts that *Arabidopsis* LysM genes and NFR-type genes have a mechanistic similarity in their perception of bacterial exudates and it is differentiation in the downstream signalling cascade which leads to basal defence reaction or nodulation formation respectively.

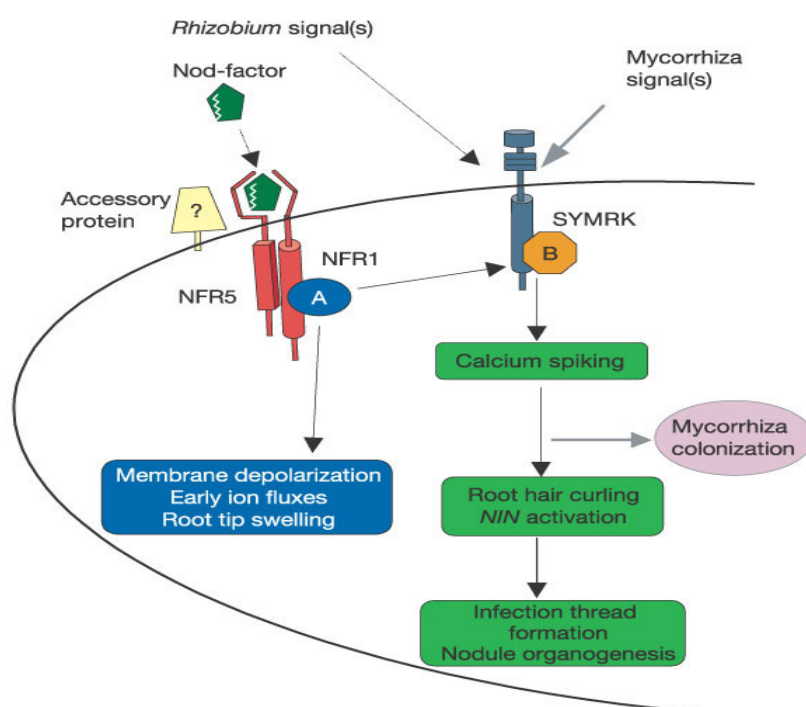


Figure 5.1 Model of the mechanisms of Nod factor function which may share some functional similarity and evolutionary origin with a LysM mediated basal resistance response. Published with permission of Nature publishing group, from the papers of (Radutoiu et al., 2003).

5.2 Assessing LysM gene function in basal resistance

The substantial natural variation observed in basal resistance across the *Arabidopsis* population probably originates from a variety of sources, for example, transcriptional differences between ecotypes (Weigel and Nordborg, 2005). Variation in the LysM gene sequence may exist between multiple ecotypes so if this trend continues, the variation may affect efficiency of pathogen perception as it did for the *RPS2* gene (Caicedo et al., 1999). In addition, the combination of LysM genes working in conjunction in a particular ecotype may differ in their signal perception efficiencies. This chapter aims to characterise the LysM gene family and find evidence for their involvement and influence in basal resistance. In addition to determining sources of variation within this family, this may explain some of the substantial natural variation in basal resistance efficacy.

Three genetic approaches have the potential to reveal the biological role of LysM genes in basal disease resistance: T-DNA insertion, RNAi and 35S over-expression. The over-expression strategy, however, was unlikely to have a significant effect on the trait if LysM genes are acting upstream in the signalling pathway or directly involved in pathogen perception (Zipfel, 2008). Preliminary work assessing the potential for a multiple RNAi approach was set aside due to problems of poor sequence homology between LysM receptor-like Kinases (Schwab et al., 2006; Ossowski et al., 2008). Knock-down efficiency was also anticipated to be a confounding factor as predicted rates within the literature estimated an average 80% reduction in gene expression (Schwab et al., 2006; Zhu et al., 2007). However this level of reduction in receptor gene expression may be compensated by resistance signalling amplification resulting in no noticeable reduction of defence response or trait variation (Schwab et al., 2006).

The T-DNA knockout approach was therefore chosen as it may provide a potential for clear results while considering the hypothetical mode of action and function of LysM genes. The selected strategy involved screening a large variety of T-DNA knockout lines of individual LysM genes with a focus on the RLK sub-set (Young et al., 2001; Ulker et al., 2008). The resulting true knockout lines were used to study the effect of individual LysM RLK knockouts on the basal resistance.

5.3 Results

5.3.1 Summary of LysM genes and annotation information

Below is a figure summarising the locations of *LysM* genes in addition to the well known PAMP genes *FLS2* and *EFR* (Fig. 5.2). The description of the characteristics of the 15 LysM genes is based on the TAIR website annotation information. Each of the LysM genes reportedly exhibit cell wall catabolism and all but two are currently predicted to be membrane bound. For 2 of the LysM domain-containing genes (AT5G08200, AT5G23130) the location is not currently predicted however, this is not unusual as annotation information is not complete for all *Arabidopsis* genes. Six of the 15 LysM genes reportedly have receptor-like kinase function and are marked on Fig 5.2 by an asterisk. For the remaining LysM genes, function is currently unknown although all LysM domains have now been associated with peptidoglycan binding.

The EFR and FLS2 products are both involved in: ATP binding, receptor-like kinase activity, membrane bound, LRR domains and serine/threonine kinase activity. These genes share some of their characteristics with members of the LysM family, in particular the subset of 6 membrane bound LysM RLK genes which also show ATP binding in the case of AT3G01840 and AT3G57120. *CERK1*, in addition to AT3G01840, AT3G57120 and AT1G51940, shares further similarity with *EFR* as they all show tyrosine kinase function which may be involved in signal transduction (Espinosa and Alfano, 2004). Two LysM RLKs, AT2G33580 and AT2G23770, show further similarity with *FLS2* as they are predicted to contain serine/threonine domains which have been previously shown to specifically phosphorylate MAP kinases which, in the case of *FLS2*, results in signal cascade transduction and activation of basal defensive compounds (Madsen et al., 2003; Qiu et al., 2008).

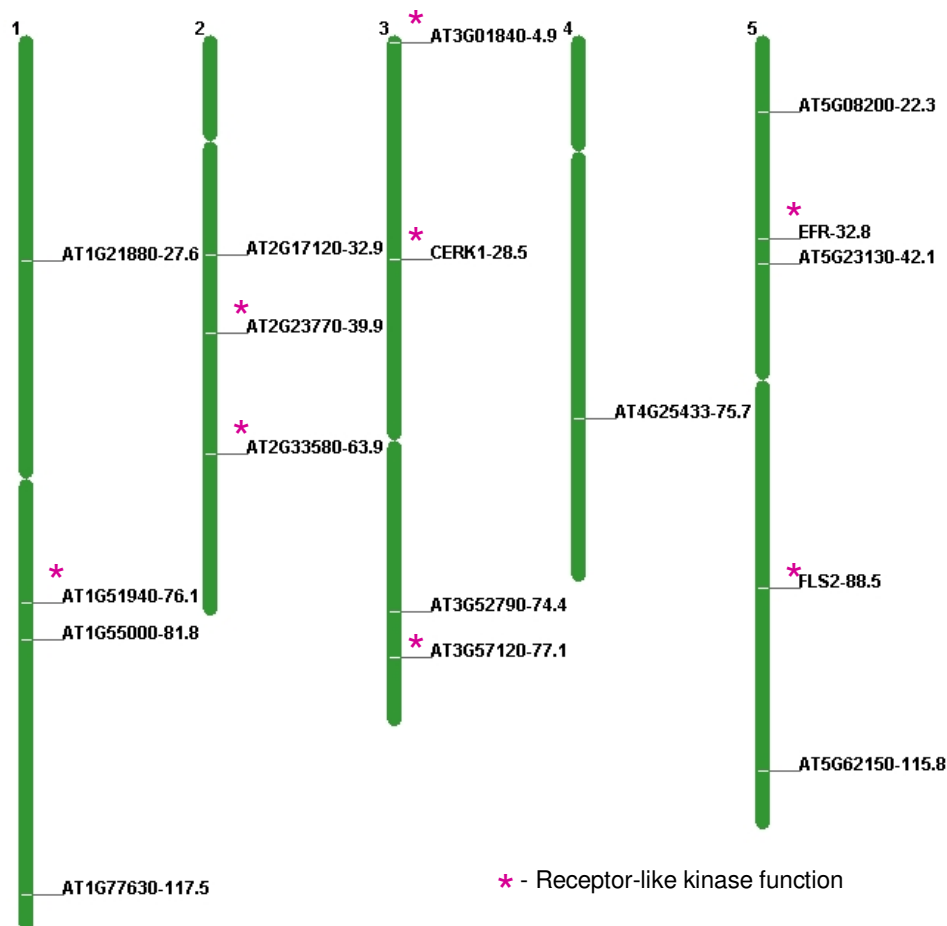


Figure 5.2 *Arabidopsis* chromosome map displaying all 15 LysM genes in addition to *EFR* and *FLS2* PAMP genes. Genetic distance, measured in cM, is annotated next to each of the genes and RLK function is denoted by an asterisk.

5.3.2 BLAST searches of LysM genes

BLAST searches between LysM genes and genes of closest nucleotide sequence homology show *CERK1* has a 71 % identity sequence match (E-value=6e115) to *NFR1* from *Lotus japonicus* (blastn comparative analysis-<http://blast.ncbi.nlm.nih.gov/Blast.cgi>). Several other LysM genes have relatively close homology to *NFR1*, as described in section 5.3.1. High LysM gene homology to *NFR1* is correlated with predicted kinase function and close association with QTL peak genetic location. However, this is not a universal trend as some LysM genes show strong affinity with QTL location and are relatively poorly related with *NFR1*.

DNA sequence comparison using BLAST shows strong sequence homology between LysM genes from *Arabidopsis* and those of other species. For example,

AT1G21880, AT1G77630 and AT2G17120 have strong homology (>70%) with a *Ricinus communis* LysM gene (GeneID:8274109) which reportedly encodes a general peptidoglycan-binding module (Bateman and Bycroft, 2000). AT5G23130 shows a 73% nucleotide homology to a LysM gene from *Zea mays* (GeneID:100283704), which as yet has no reported function. DNA sequence identity exists between the *Arabidopsis* LysM gene AT2G23770 (84% identity) and AT2G33580 (64% identity) with *Sesbania rostrata* LysM-RLK *SrLyr3* gene which is significantly expressed during nodulation formation. In addition, *SrLyr3* shows 85% similarity to the DNA sequence of *MtLyr3* from *M. truncatula*, which is involved in nodulation formation (Capoen et al., 2007).

5.3.3 LysM gene phylogenetic trees

A phylogenetic tree was generated showing relative homology of nucleotide sequences within the LysM gene family (Figure 5.3). The tree image also highlights the 6 LysM genes (marked with asterisk) with receptor-like kinase function and their increased number in proximity to QTL peaks compared to other LysM genes. The numbers at each branching point of the image represents the bootstrap analysis confidence level based on Felsenstein's bootstrap method that: if the tree were generated a 100 times, that the same configuration of tree would be produced; the closer to 100 the number, the more confident the prediction (Efron et al., 1996). The accompanying annotation of the phylogenetic tree describes the bacterial strains and RIL populations, where QTLs coinciding with LysM genes have been found. The results show a subset of RLK LysM genes that are associated with more of the generated QTLs than the remaining LysM genes without RLK function.

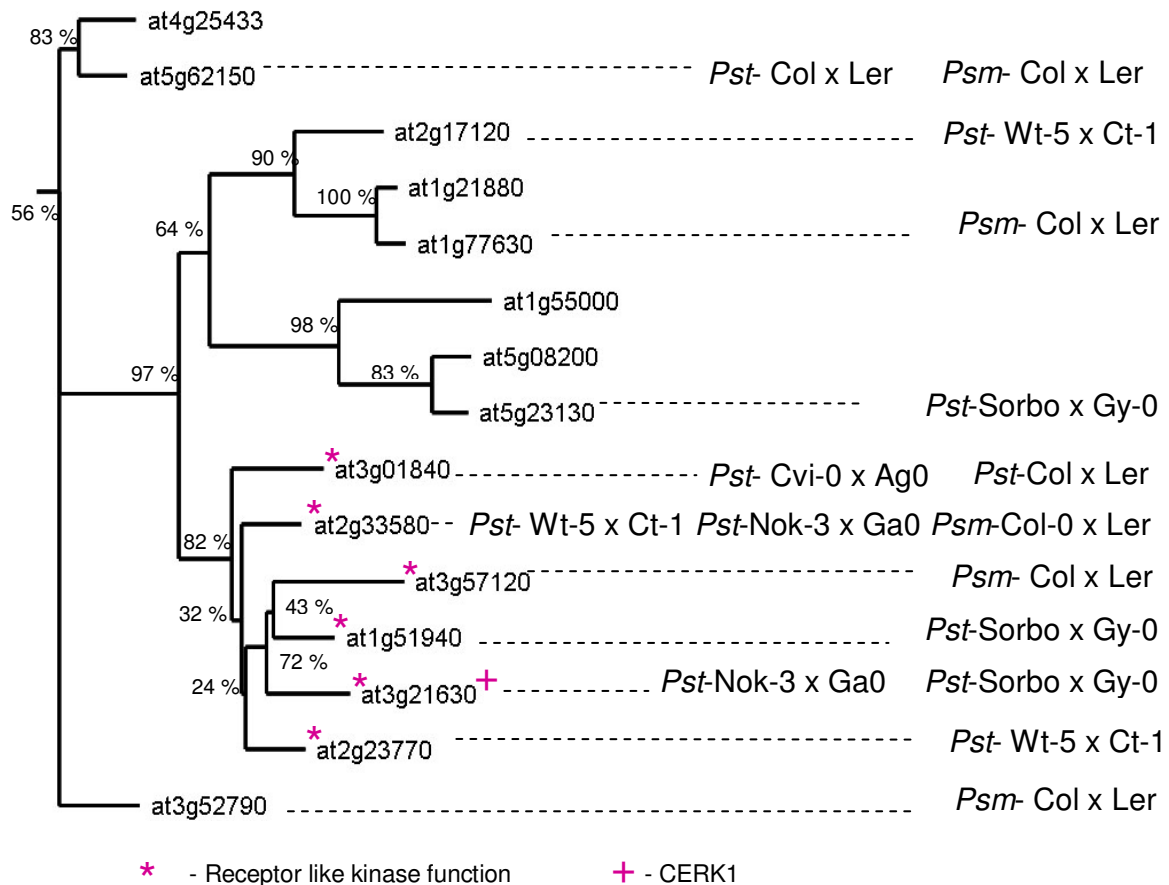


Figure 5.3 Phylogenetic tree of the nucleotide sequence of the LysM genes. The tree illustrates the relative homology of LysM genes and their correlation to QTLs identified by bacterial growth assays on different RIL lines following infiltration with *Pseudomonas* strains. Numbers at each branching point of the image represent the bootstrap analysis confidence level based on Felsenstein's bootstrap method, 100 bootstrap replications.

5.4 Expression profiles of LysM gene family

5.4.1 Developmental and anatomical profiling

The heat map output of normalized signal intensity values images (Fig. 5.4 and 5.5) represents the absolute expression value as a colour scaled in relation to the expression potential of each gene. Therefore, the darkest colour equates to the highest expression value of an expression vector (gene expressed in given parameter). Bioinformatic analysis has verified the output from Genvestigator in comparison with the raw ATH1 (developed by Affymetrix) full genome array data and provides interpretation information (for example see (Zimmermann et al., 2004). Developmental meta-profiling shows constant low expression rates across

developmental stages for all LysM family genes, which is consistent with the profile expected from a receptor family (Fig.5.4). The developmental expression pattern shows no definitive pattern across the LysM genes, which would be indicated on the profile by an expression trend which runs horizontally across a particular growth stage. This is consistent with the known information for currently identified disease receptors which are generally expressed consistently at a low level throughout the lifecycle of a plant.

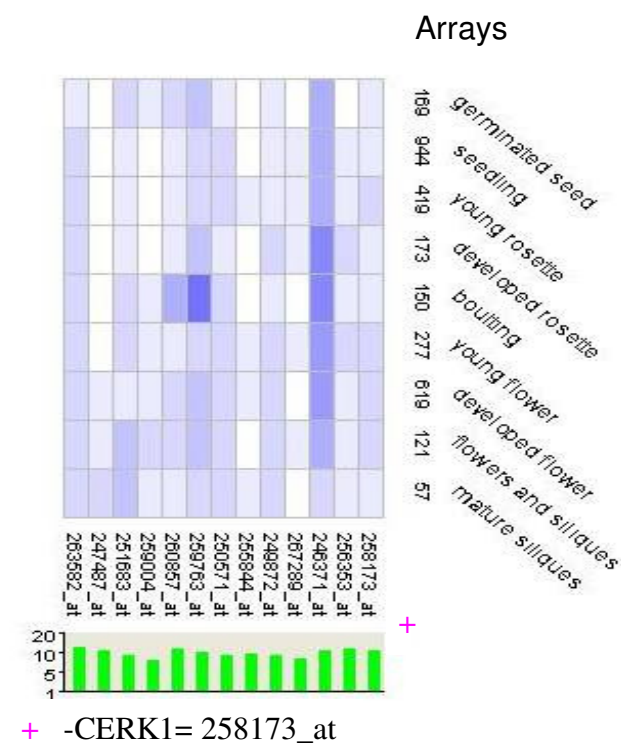


Figure 5.4 Genevestigator developmental expression profile for the LysM gene family and the number of arrays which make up the pooled data. The darker the heat map colour the greater the expression potential of a particular expression vector. White indicates no change between conditions. The green graph below shows the number of arrays data is pooled from. LysM genes listed from left to right: AT2G17120, AT5G62150, AT3G57120, AT3G01840, AT1G21880, AT1G77630, AT5G08200, AT2G33580, AT5G23130, AT2G23770, AT1G51940, AT1G55000, AT3G21630.

Figure 5.5 illustrates the anatomical expression profile for the LysM family and shows broad similarity between the LysM gene expression with low to moderate levels of expression across all anatomical categories. Comparisons of LysM gene anatomical expression category separation shows no discernable pattern between categories (e.g. high expression in roots and little expression in leaves) or between LysM genes for any of the tested genes. This shows a general similarity for LysM genes to produce a relatively low expression across plant developmental and

anatomical categories. This is consistent with other pathogen receptors such as FLS2, although it should be noted there are some areas of relatively high expression compared against their own internal expression ratios (Zipfel et al., 2004). For example, relatively high expression values were observed for both AT5G08200 (250571_at) and AT5G23130 (249872_at) in pollen or AT2G33580 (255844_at) and AT2G33580 (259004_at) in the lateral root cap when compared against a gene's own average expression values. LysM genes are differentially expressed at the defined anatomical categories and there is no discernable expression relationship found or explained by phylogenetic sequence similarity between LysM genes.

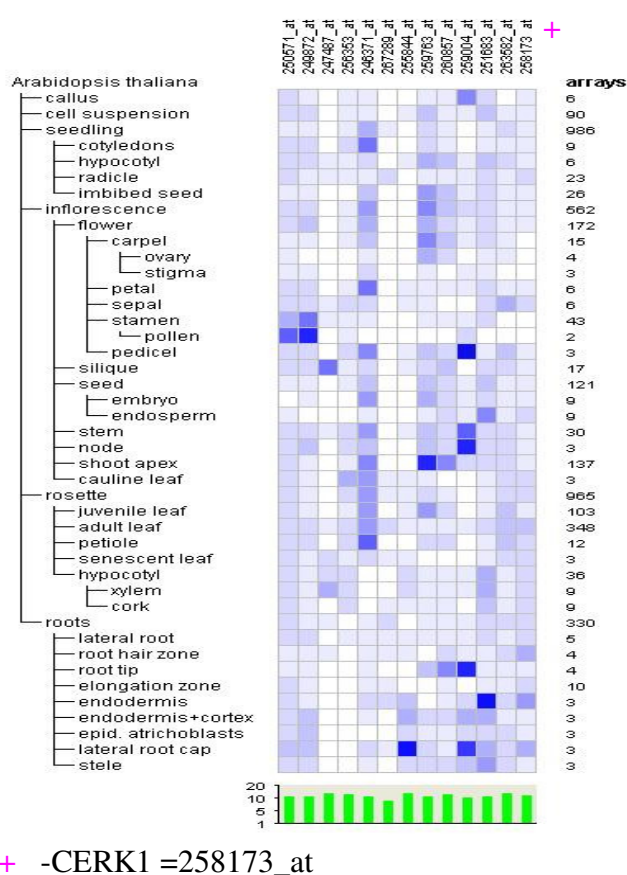


Figure 5.5 Genevestigator anatomical expression profile for the LysM gene family and the number of arrays which make up the pooled data. The darker the heat map colour the greater the expression potential of a particular expression vector. White indicates no change between conditions. The green graph below shows the number of arrays data is pooled from. LysM genes listed from left to right: AT5G08200, AT5G23130, AT5G62150, AT1G55000, AT1G51940, AT2G23770, AT2G33580, AT1G77630, AT1G21880, AT3G01840, AT3G57120, AT2G17120 and AT3G21630.

5.4.2 Stimuli induced expression profiling

Figure 5.6 shows heat map interpretations of various biological, chemical and environmental stimuli and its effect on LysM gene expression. Several points of interest have been highlighted with red arrows including the similarity of increased expression for AT3G21630 (*CERK1*) LysM gene following chitin and *P.syringae* treatments. The heat map indicates a down regulation of LysM when exposed to heat and osmotic stress, suggesting a possible antagonistic relationship similar to those observed for other defence related genes. However, the profile indicates possible down-regulation of multiple LysM genes in response to abscisic acid (ABA), which is typically associated with abscission mediation, stomatal closure and suppression of plant defence responses (Fan et al., 2009). Treatment with SA, which is typically associated with the SAR response, shows 4 LysM genes up-regulated after treatment. Treatment with bacterial *EF-Tu* gene product, which is a known PAMP typically involved with binding to the *Arabidopsis* EFR receptor, induces substantial up regulation of several LysM genes, including *CERK1*. Expression profiling data is limited as the results are pooled from many hundreds of microarray experiments with different test parameters. Genevestigator heat maps are an aggregate of experiments that may not have been specifically conducted to test a particular stimulus; as a result, only inferences can be made without qPCR (or equivalent) validation. In Fig. 5.6 LysM genes listed from left to right: AT3G21630, AT1G55000, AT1G51940, AT2G23770, AT5G23130, AT2G33580, AT5G08200, AT1G77630, AT1G21880, AT3G01840, AT3G57120, AT5G62150, AT2G17120.

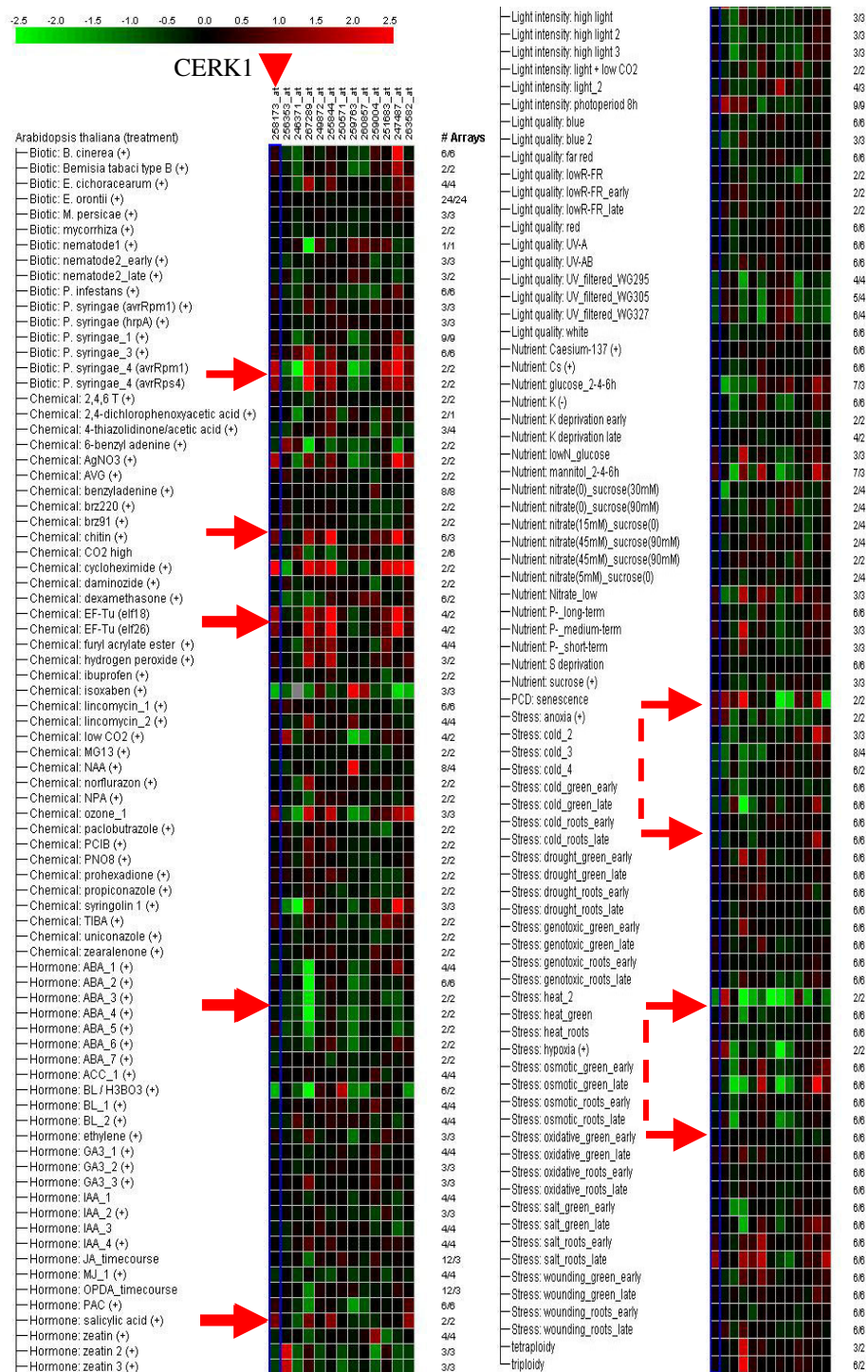


Figure 5.6 Geneinvestigator stimuli heat map displays the results for the LysM family of genes, with colour coding representing relative expression values. Columns represent probe sets, and rows represent stimuli. All gene expression profiles were normalized for colour from red through black to green. Red and green indicate relatively higher and lower expression levels, respectively, upon stimulation. Black indicates no change between conditions. The red arrows indicate stress stimuli inducing LysM gene expression of notable interest as described above (section 5.4.2).

5.4.3 Transcriptional changes in response to flg22 extract

Using microarray data, the transcriptional changes for LysM genes were monitored before and after exposure to flg22 peptide (Zipfel et al., 2004). The results show the LysM gene AT2G33580 up-regulated by 13.3 fold, AT2G23770 by 6.7 fold and a 2.4 fold increase for the *CERK1* gene. The expression of these three LysM genes is significantly increased but are comparable to other PAMP receptors (e.g. FLS2 -3.5 fold, EFR -10.0 fold), presumably as transcription increases to enhance plant sensitivity to pathogens after initial exposure (Zipfel et al., 2004). However, receptors must already be in place to detect initial pathogen exposure therefore, large fluctuation in transcription level after initial contact would be unusual for both PAMP and *R* gene defence receptors (e.g. RPS2 -4.3 fold, RPM1 -3.7 fold) (Zipfel et al., 2004). Expression changes for other LysM RLKs were not statistically significant at the time points examined in this study.

5.5 LysM gene sequencing

Deduced protein sequence data for the *CERK1* LysM gene was obtained for various *Arabidopsis* genotypes of interest from the D.Weigel group (Max Planck, Germany). Several polymorphisms were detected within the protein sequence. However, when compared to basal resistance trait data, no direct correlation could be distinguished between basal resistance and protein polymorphisms.

Nucleotide sequences and the deduced protein data were obtained for *CERK1* from different parental accessions in collaboration with V. Lipka group (JIC, UK). Significant polymorphisms were detected between genotypes which may alter the tertiary structure of the proteins. A change in tertiary structure could lead to a difference in protein function, but no direct correlation between trait data and sequence polymorphism could be accurately defined.

Nucleotide sequence for each of the 15 LysM genes was obtained for the Col-0 and Ler genotypes from TAIR and Monsanto. Sequence comparison showed a small number of polymorphisms in 4/15 LysM genes and no direct correlation between trait data and polymorphisms.

5.6 T-DNA insertion knockouts

Assessment of basal disease resistance in *Arabidopsis* lines carrying T-DNA insertion within the LysM gene family aimed to determine whether these genes are required for basal disease resistance. Twenty-eight T-DNA lines were selected which targeted all of the 6 LysM-RLKs as well as several other LysM members without the kinase domain. The experiment was set up as a double-blind design where plants of each line were tested for bacterial growth before the homozygosity of T-DNA insertion was determined. Control plants of both Col-0 (the background of the T-DNA insertions) and Ler were included in every experiment and positioned for comparative reference during the analysis. Statistical analysis was used to establish any significant difference in resistance/ susceptibility between wild type plants and LysM knockout lines.

5.6.1 Trait assessment of T-DNA lines

Sixteen plants from each of the segregating T-DNA insertion lines were inoculated with an OD₆₀₀ 0.002 and luminescence of 3 leaf discs per plant was measured 48 hrs post infiltration. Four plants of Col-0 and Ler were placed in each tray as controls. Data of bacterial growth was pooled and incorporated into a comprehensive graph (Figure 6.6) where bacterial growth in plants carrying heterozygous or homozygous T-DNA insertions were compared with wild type controls.

5.6.2 PCR assessment of T-DNA insertion type

Gene specific primers, LP and RP, were used with T-DNA left border primers (LBb1.3) to assess each of the T-DNA lines. Products of LP to RP (no insertions-wild type) were designed to generate products of around 900-1100 bp. Insertion in both chromosomes (homozygous) will result in a product from RP to insertion site 300 + N bases, plus the 110 bases from the insertion LBb1.3 to the left border of the vector (<http://signal.salk.edu/tdnaprimers.2.html>). Heterozygous plants generate bands of both sizes.

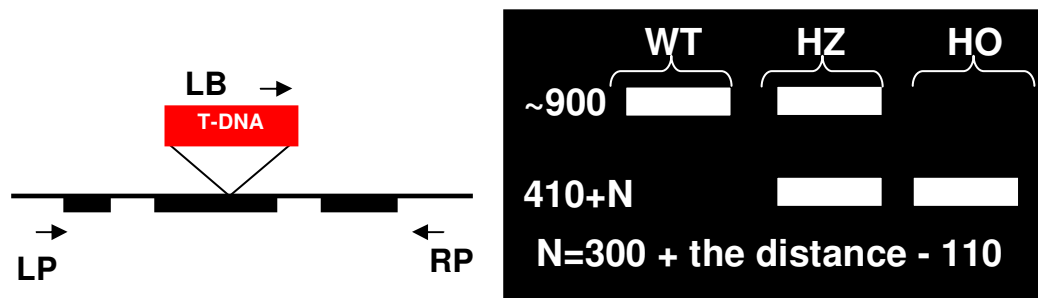


Figure 5.7 Illustration of T-DNA insertion site within a gene relative to primer target location (left). Example shows two paired reactions to differentiate between insertion types (right). Images based on Salk institute images (<http://signal.salk.edu/tdnaprimers.2.html>).

PCR products were resolved with multiple gels and only a sample annotated image is presented below. The image was selected to show the variety of T-DNA insert genotypes found in the lines, however similar patterns are seen across all lines.

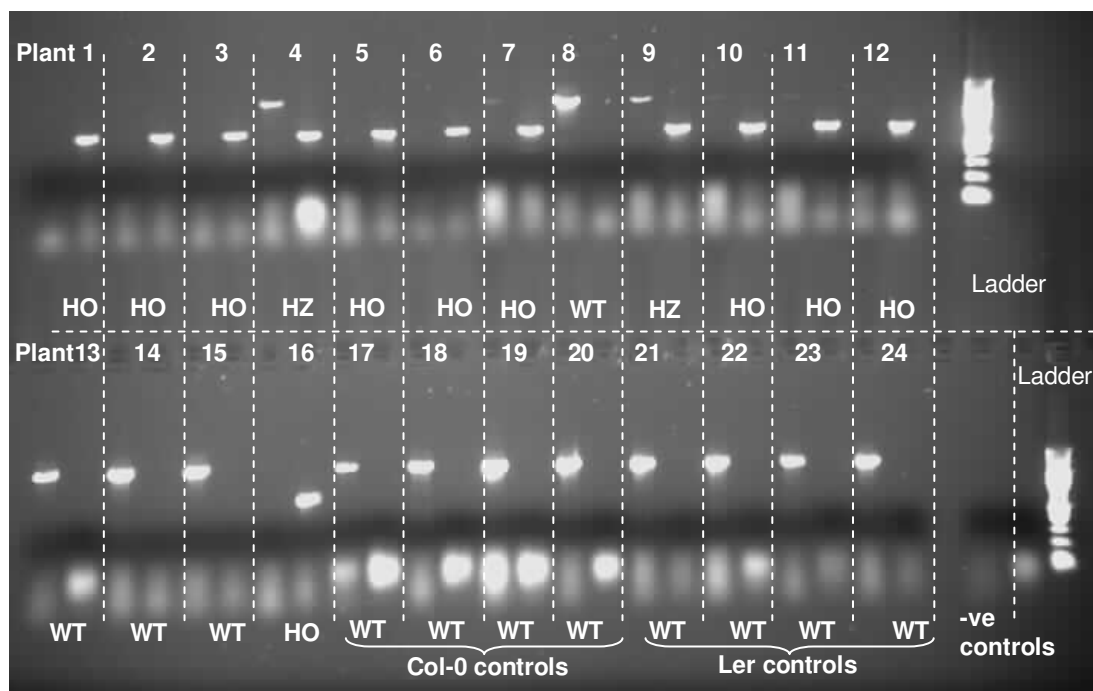


Figure 5.8 Agarose gel image of the PCR product from T-DNA lines and controls for genotype assessment of the T-DNA insertion. Salk_102100 line knockout in the AT3G57120 gene.

5.6.3 Combined trait data and PCR insertion assessment

A selection of 28 T-DNA insertion lines were surveyed for their basal resistance and then tested to establish the T-DNA genotype of the lines. From this data, 14 lines were discounted because the segregation data showed only wild type plants or genotype assessment was inconclusive. The results from the rest of the 14 lines are shown below, and separated to show individual comparison with Col-0 controls and heterozygous T-DNA inserts where present. Assessment of the majority of homozygous T-DNA insertion lines show significant differences in bacterial growth compared to Col-0 control, where single LysM genes are knocked out. There are some notable exceptions to this trend. It should also be noted where these T-DNA insertions are located in relation to the LysM gene of interest, as this may affect the likelihood of effective gene transcription reduction.

To cross-compare between T-DNA insertion lines, Fig. 5.9 displays *Pst* assay bacterial growth data as a calculated percentage of heterozygous or homozygous T-DNA insertion lines, compared against internal Col-0 controls within each experimental line. Comparison against internal controls provides a more comparable standard to assess relative differences between lines and between homozygous and heterozygous insertions. Data shows significant differences in bacterial growth between lines targeting the same and different LysM genes as well as significant differences between homozygous and heterozygous insertions (additional data Table 5.1).

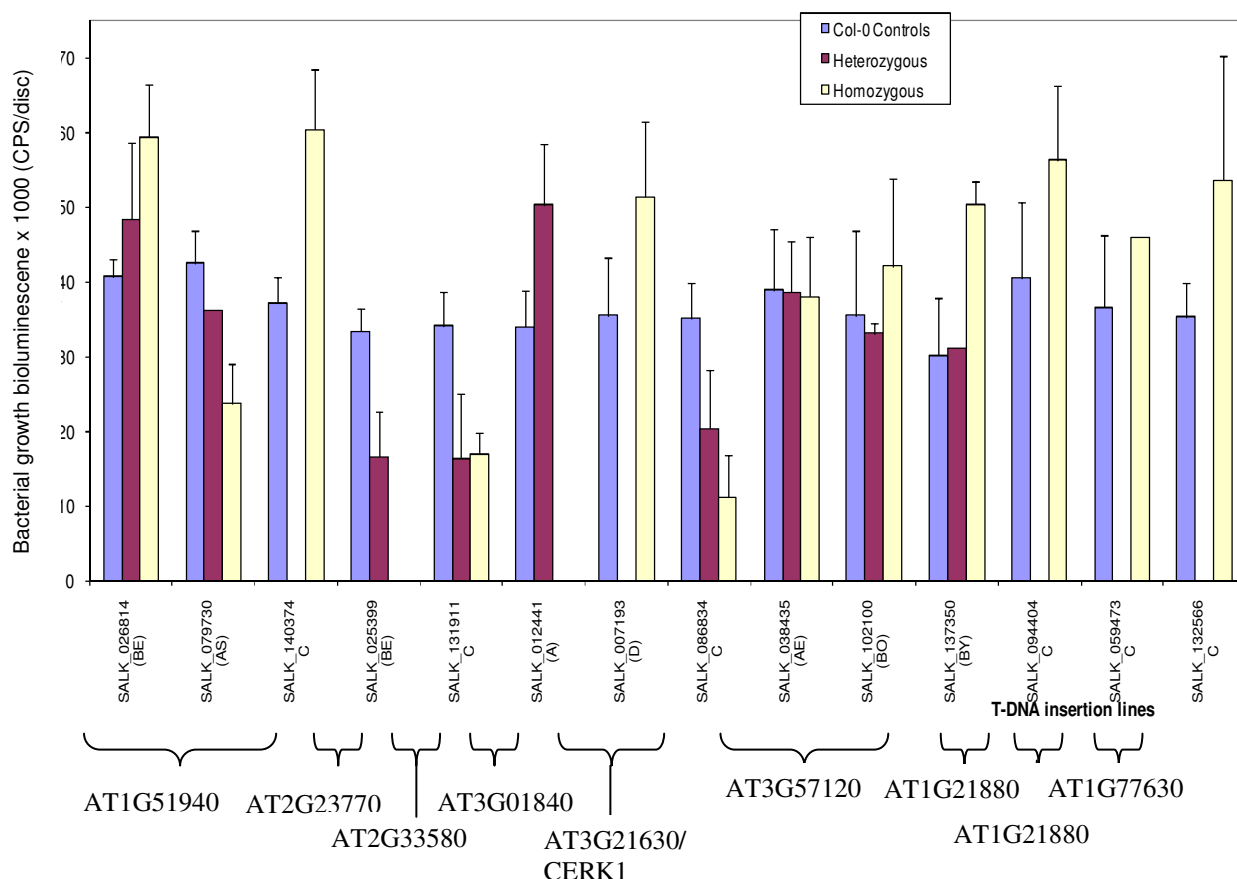


Figure 5.9 Bacterial growth data is displayed as Lux light emission from the tagged *Pst* line for heterozygous and homozygous T-DNA insertion plants compared against internal Col-0 controls within each experiment. The error bars represent standard deviation (SD). Below the graph the knockout lines are grouped according to which LysM genes are targeted. Homozygous and heterozygous insertion lines with significant difference ($P = <0.05$) in bacterial growth compared to wild type controls are described in the additional statistical analysis data in Table 5.1.

‘Genstat’ software was used to analyse the Lux infiltration trait data from knockout lines and, as expected, found statistical differences between knockout lines and between insertion genotypes when compared as a whole. ANOVA analysis was used to individually compare the trait data from heterozygous and homozygous insertion knockout genotypes against the internal Col-0 control. Several significant differences were found, showing differences in bacterial growth scores when a LysM gene is knocked out compared against internal Col-0 controls; the probability results are summarised (Table 5.1).

LysM gene target	Knock out line	Heterozygous insertion genotype vs. Col-0 control P value	Sample No. plants	Homozygous insertion genotype vs. Col-0 control P value	Sample No. plants	Insertion site
AT1G51940	SALK_026814 (BE)	0.174	9	0.028 **	2	Intron
AT1G51940	SALK_079730 (AS)	0.484	2	0.001 *****	7	Intron
AT1G51940	SALK_140374 C	N/A	0	0.001 *****	15	Exon
AT2G23770	SALK_025399 (BE)	0.003 ****	12	N/A	0	300 UTR5
AT2G33580	SALK_131911 C	0.001 *****	3	0.001 *****	5	Exon
AT3G01840	SALK_012441 (A)	0.001 *****	14	N/A	0	Exon
AT3G21630	SALK_007193 (D)	N/A	0	0.010 **	4	300 UTR5
AT3G21630	SALK_086834 C	0.007 ****	4	0.001 *****	2	1000-Promotor
AT3G57120	SALK_038435 (AE)	0.970	4	0.911	6	300 UTR3
AT3G57120	SALK_102100 (BO)	0.781	2	0.272	10	Exon
AT3G57120	SALK_137350 (BY)	0.941	2	0.065*	2	Exon
AT1G21880	SALK_094404 C	N/A	0	0.011 **	16	1000-Promotor
AT5G08200	SALK_059473 C	N/A	0	Insufficient data points	1	Intron
AT1G77630	SALK_132566 C	N/A	0	0.032 **	15	Exon

Table 5.1 Table summarising analysis of variance statistics comparing the trait data from heterozygous and homozygous insertion genotypes against the Col-0 controls. Insertion site for the T-DNA and the LysM knockout target data is also included for comparison. Sample number is included for insertion lines, however this number is variable as most lines are still segregating for the insertion and so tested as heterozygous, homozygous or wild type. Significance value – *0.1 **0.05 ***0.01 ****0.005 *****0.001

5.7 Discussion

5.7.1 Phylogenetic analysis and LysM gene properties

Well known PRRs involved in basal resistance such FLS2 and EFR are predicted to share some common structural properties which aid in their ability to act as pathogen receptors such as; ATP binding, being membrane bound, transmembrane serine/threonine receptor kinase activity and LRRs. These properties are predicted in other PAMP receptors, such as CERK1, a fungal PAMP receptor, which also displays transmembrane serine/threonine receptor kinase activity with an ATP binding site, however it lacks LRR but has predicted peptidoglycan binding function (section 5.3.1). The prediction is that the majority of LysM genes are membrane bound, are involved in amino acid phosphorylation and involved in cell wall catabolism, serine/threonine kinase activity. However, not all LysM genes have the same annotation characteristics and there is notable variation amongst this gene family in predicted receptor-like kinase function. Recent evidence and updates to the annotation information have linked the family of genes containing a LysM domain to peptidoglycan binding. The duality of the CERK1 chitin binding and peptidoglycan binding may be resolved by a conserved structural component, common between the two cell wall structures such as N-acetylglucosamine (GlcNAc) (Miya et al., 2007).

Phylogenetic comparison of the nucleotide and protein coding sequences between each of the LysM family shows that genes with receptor-like kinase functional domains are more closely related, essentially forming a small subset (Fig. 5.3). BLAST searches and phylogenetic comparisons show significant similarity between LysM genes such as *CERK1* and the *NFR1* gene from *Lotus japonicus*. Root nodulation in leguminous plants for symbiotic nitrogen fixation has been linked to the *NFR1* protein, which has been postulated to act as a Nod factor receptor itself or involved in the earliest generation of the Nod factor signalling cascade (Radutoiu et al., 2003; Arrighi et al., 2006). Nucleotide based phylogenetic tree analysis indicates homology to *NFR1* is correlated with kinase function and this in turn correlates with significant QTLs however, some LysM genes show strong affinity with QTL location but show relatively weak homology with *NFR1*. The phylogenetic trees show LysM genes with RLK functional domains grouping into a subset based on code similarity. This subset has a higher association with QTL peaks that correspond to LysM gene

locations compared with other LysM genes in the family. Therefore attention should be focused on this small subset of LysM RLK genes in order to determine a conclusive link to basal disease resistance.

5.7.2 Transcriptional analysis

Transcriptional analysis of the LysM genes following exposure to an array of biotic and abiotic stimuli shows that, out of 15 lysM genes, only two fail to show strong gene expression changes when exposed to *P.syringae*. The LysM gene *CERK1*, which reportedly perceives chitin oligosaccharides from fungi, shows increased expression rates when exposed to chitin (Wan et al., 2008). However, expression rates are similarly increased when exposed to *P.syringae*, possibly suggesting that it may not be exclusively a fungal elicitor interaction with CERK1, and bacterial exudates may also be involved (Miya et al., 2007).

Several LysM genes are up-regulated following treatment with the *EF-Tu* gene product (elf18), a known PAMP and bacterial elongation factor, which is typically associated with binding to the *Arabidopsis* *EFR* receptor and so instigating a basal resistance defence reaction (Nurnberger and Kemmerling, 2006). This up-regulation observation may suggest possible interaction or a general up-regulation of PAMP genes (including LysM genes) following elf18 exposure to heighten pathogen sensitivity as occurs with a SAR response (Glazebrook et al., 2006).

When exposed to heat stress and possibly osmotic stress, the majority of LysM genes are significantly down-regulated which, if LysM were proven to be involved in the disease resistance pathway, would demonstrate a redirection of metabolic resources away from disease related genes when a plant is under stress. Redirection of finite metabolic resources away from disease or wounding related genes in response to abiotic stresses, such as heat shock, has previously been shown (Cheong et al., 2002). Therefore, as LysM genes are strongly down regulated following abiotic stress exposure, and are up-regulated by elicitor exposure, this is consistent with the hypothesis that LysM genes are involved in disease resistance.

Genevestigator's expression map shows several LysM genes are up-regulated in their transcription rates after exposure to salicylic acid (SA), which has been consistently linked as a signal for systemic acquired resistance (SAR) (Ryals et al.,

1996). LysM genes could be stimulated in preparation to defend from secondary infection as an enhancement of existing defences commonly associated with SAR (Glazebrook et al., 2006).

It is possible to examine specific microarray data, derived from gene chips which indicate transcription level, selecting experiments which closely matched the criteria applied to this study in an experiment monitoring gene transcription after exposure to flg22 peptide (Zipfel et al., 2004). Three LysM genes (AT2G33580, AT2G23770, AT3G21630) show changes which are significant (13.3, 6.7, 2.4 fold increases respectively), but are consistent with the transcription increases to enhance plant sensitivity after initial exposure as other pathogen receptors such as FLS2, EFR and RPS2 (section 5.4.3) (Chen and Chen, 2002; Xu et al., 2006). Typically, pathogen recognition receptors would be expected to be expressed prior to disease exposure. Following initial detection of the invading pathogens, the transcription of genes involved in signal amplification and disease responses are enhanced. Initial perception receptors would typically not be up-regulated, except perhaps to prime plant tissue resulting in improved sensitivity for additional pathogenic challenge. Expression changes for other LysM RLKs were not statistically significant at the time points examined in this study.

Transcription level is an indication of trait modification but not conclusive as many other factors are influential, such as post-translational modification and the regulatory influence of rate-limiting enzymes or the affect metabolic intermediates have on the final gene action (Fell and Thomas, 1995). For example, post translational redox state regulates the conformation of the *NPR1* gene, involved in SA-dependent defence gene expression (Despres et al., 2003). *NPR1* links pathogen-triggered redox changes in plant tissue with defence gene regulation as it acts as a co-factor, dependent on the oxidation state of cystine residues in its structure to mediate the interaction with TGA transcription factors (Tada et al., 2008; Spoel et al., 2009).

5.7.3 Sequencing analysis

For the protein sequence data and the nucleotide sequence, multiple substantial polymorphisms were detected between genotypes. These variations in sequence have the potential to alter the tertiary structure and so may alter the effectiveness or

the binding ability of LysM genes, which may account for some of the large trait variation observed in the *Arabidopsis* population (Schmid et al., 2003a). Unfortunately, a direct correlation between any specific polymorphism and alterations in trait data could not be established. There is a possible trend which suggests increasing numbers of polymorphisms are associated with increased disease susceptibility, however the genotype sample size is not large enough to define this. There are only small variations found between LysM genes when complete coding sequence of Col-0 and Ler are compared, however the disease trait scores are fairly similar so significant sequence differences were not expected. Overall, there are multiple synonymous and non-synonymous polymorphisms found in the LysM genes which would suggest that the family is still in evolutionary flux (Ellis et al., 2000). This is consistent with other disease related genes which are typically located within an evolutionarily adaptive chromosomal area to match plant specificity with pathogenic alterations (Mauricio et al., 2003; Jones and Dangl, 2006). The variation observed in the sequence of LysM genes is unlikely to be a critical factor affecting basal disease resistance in the majority of cases. However, polymorphism may be a contributing factor that could be defined further if direct proof of LysM gene family involvement can be established.

5.7.4 T-DNA knockout lines

Previous phylogenetic analysis and cross comparison with QTL peak locations indicated that 6 LysM genes with RLK functionality were most likely to be influential in basal resistance therefore knockout experiments were focused on these. Other LysM gene knockouts were tested but only where they appeared in the exon as it is difficult to predict whether insertion in the promoter or intron region would affect expression. Twenty-eight knockout lines derived from Col-0 lines were chosen covering 9 separate LysM genes, with 10 of the lines reportedly homozygous for the T-DNA insertion. Eighteen knockout lines were still segregating at various ratios. Where possible, inserts were located within exons (17 out of 28). Insertions were located in the promoter in 7 lines, and 4 out of the 28 had insertions within an intron.

The knockout line testing was designed as a double-blind experiment where the bacterial Lux count trait data for each plant was unknown as well as their insertion

genotype (wild-type, heterozygous, homozygous). Bacterial trait data was collected first from the 28 knockout lines obtained by infiltration with Lux tagged *Pst*. PCR genotyping was then conducted to establish the type of insertion an individual plant has e.g. wild-type, heterozygous or homozygous T-DNA insertion (section 5.6.2). Fig. 5.9 shows the bacterial trait data for 14 lines of plants, that the PCR genotyping showed to be suitable for this experiment, separated into insertion types (homozygous and heterozygous) and compared against Col-0 controls.

The majority of knockout *Arabidopsis* lines show increased susceptibility in comparison with Col-0 controls after knockout of LysM genes, although some lines show an increase in resistance so this is not a definitive trend across all LysM knockouts. Significant differences in Lux trait data were observed in 10 of the knockout insertion lines tested, 8 significant differences were found in homozygous T-DNA insertion genotypes and 4 in heterozygous insertion genotypes (Table 5.1). A line with low homozygous plant replication (Salk_137350) had a probability approaching 95% significance ($P=0.065$). Another line (Salk_059473) produced a single homozygous plant, which prevented ANOVA analysis but T-test showed to also approach significance ($P= <0.1$), which may show significant deviation from the controls with further replication. Compared with controls, knockout lines typically showed differences of approximately 40-60%, which is modest, but does indicate direct evidence of LysM gene involvement in basal disease resistance trait. The differences observed are not limited to LysM RLK genes which may suggest broad LysM gene family involvement, however the mechanisms of this can only be hypothesised.

When comparable pathogen receptors are knocked out, some insertion line experiments have shown higher effects on trait. Experiments such as the T-DNA knockout of MEKK1, a mitogen-activated protein kinase involved in the activation of the map kinases signalling cascade, completely prevents plants from activating MPK4 in response to flg22 treatment, resulting in an approximate 1 log increase susceptibility (Suarez-Rodriguez et al., 2007). The loss-of-function mutation of the FLS2 receptor, *fls2-17*, showed increases in bacterial susceptibility of ~0.7 log compared to wild-type Col-0 when sprayed with an already virulent pathogen, *Pst* DC3000 (Zipfel et al., 2004). Double T-DNA insertions in the PUB17 line has produced knockouts of the R-genes *RPM1* and *RPS4* and subsequently compromises their resistance against *Pst* containing the corresponding Avr genes and increase

susceptibility by approximately 10 fold (Yang et al., 2006). When critical components of nodule formation in *Lotus japonicas*, *LjNFR1* and *LjNFR5*, are both knocked out, this eliminates almost all downstream Nod factor-inducible responses and blocks infection thread formation (Smit et al., 2007). In comparison with other published T-DNA knockout data these examples show a greater increase in susceptibility than the increase shown following knockout of single LysM genes however, it is hypothesised that LysM receptors may work in conjunction therefore, multiple knockouts maybe required to observe similar (e.g. 10-fold) susceptibility changes. *M. truncatula hcl-1* mutant line contains a single knockout of the LYK3 gene, an ortholog of the Nod factor receptor NFR1 found in *Lotus japonicas*, results the induction of ~50% less Nod factor signalling genes typically associated with this pathogen interaction (Li et al., 2010). This single gene LysM ortholog knockout would be more closely comparable to the results observed in the single *Arabidopsis* LysM knockout experiments.

Whether knockout lines show an increase or decrease in susceptibility, the homozygous knockout insertion lines have predominantly more trait difference than heterozygous lines. Where both heterozygous and homozygous insertions are represented in the lines, such as SALK_026814 and SALK_086834, they predominantly show altered bacterial growth values, falling between controls and the homozygous values. This would add weight to a hypothesis that for those LysM genes, a moderate gene dose effect may be influential as transcription rate may be as reduced in heterozygous lines. Such gene dose effects have been observed in comparisons between heterozygous and homozygous knockout insertions within Lysine synthesis genes. These comparisons showed a greater trait effect in the homozygous insertion but the heterozygous insertion was still significantly altered in comparison with the wild-type (Zhu and Galili, 2003). In the line SALK_13191C, heterozygous and homozygous insertions in the LysM gene produce similar results that are both significantly different from the control. This would indicate that the targeted gene is having a significant effect on the bacterial growth trait and that only a partial knockout is enough to disrupt its role in basal resistance. The potential hypothesis for this observation is described as haplo-insufficiency, where a single gene copy is not sufficient to maintain a trait. Similar results have been observed where homozygous and heterozygous knockout mutants in *VAMP 721/722* result in haplo-insufficient non-host and basal resistance in response to the powdery mildew

Erysiphe pisi (Lipka et al., 2010). Analysis of the RNA transcription level of the genes would be required to further elucidate the effect gene dose has on the trait for both of the observations of SALK_026814 and SALK_13191C. These lines could be tested with qPCR however there was insufficient time to verify this hypothesis in this project. Other factors such as post translational modification, redox induction of gene product activity or many other regulatory mechanisms may be at work and require further investigation. Many lines were still segregating at a greater than reported ratio which has reduced the number of replicates within the tested sample to lower than expected levels, as described in Table 5.2. Further trials should take the lower than publicised segregation ratios into account and include more replicates to confirm findings or test other knockout lines targeting different LysM genes.

5.7.5 LysM gene effect

With each of the knockout lines it is important to note the site of insertion as exon insertions typically have a higher probability of transcription disruption. In contrast, intron and promoter region inserts may have more unpredictable and potentially variable results as transcription disruption is dependent on disrupting promoter regions upstream of the gene of interest (Krysan et al., 1999; Ulker et al., 2008). Description of insertion type is listed in Table 5.1 and there is a trend towards exon insert lines showing higher susceptibility and greater differences from the controls. Similar trends can be seen when both the SALK_102100 and SALK_137350 lines target the AT3G57120 gene in the exon region. In this example both homozygous insertion lines produce a substantial, but not quite significant, increase in bacterial growth compared to controls. However, this is far from conclusive with some clear exceptions. Insertion sites appear to be an important influential factor for altering basal resistance. There is an example in the knockout of the AT1G51940 gene where two insertions (SALK_026814 and SALK_079730), both in intron regions within the gene, create a significant difference in bacterial growth between homozygous lines ($P < 0.05$). When a T-DNA insertion is not in an exon, the effect on expression levels and gene function is unpredictable (Krysan et al., 1999; Ulker et al., 2008).

The observed differences in basal resistance are moderate (approx 40-60%) in comparison with 10-fold susceptibility reductions observed in some *R* gene-mediated

responses (Aarts et al., 1998). This may be the extent of basal resistance however, this would contrast with the 50-fold differences observed in the RIL line trials. The working hypothesis of the mechanism of action suggests collaborative function, therefore cumulative knockouts of LysM genes may produce greater alterations of basal disease resistance.

5.7.6 Hypothesis of LysM gene family involvement in basal resistance

Knockout of the *CERK1* gene affects susceptibility by 40-60% in Col-0 which equates to ~20,000 CPS/disc or a difference of over 10% if compared against the range of variation observed across the selected ecotypes. Single gene knockouts of other LysM genes show similar results and suggest that various members of LysM gene family may have a role in basal resistance. Verification of expression rates is required for a definitive conclusion. However, there is sufficient evidence to propose a role for LysM gene involvement in basal resistance. LysM genes and most specifically LysM RLK genes, are postulated to be directly involved in PAMP perception. The current hypothesis is similar to the working hypothesis described previously (section 5.1), that is, suspected strong functional and mechanistic affinity with the Lotus *NFR1* and *NFR5* LysM gene binding to a lipo-chito-oligosaccharide derived bacterial signal (Madsen et al., 2003). *Arabidopsis* LysM genes show strong homology to *NFR1* and *NFR5* and have also been shown to bind to peptidoglycans (Buist et al., 2008).

The LysM *CERK1* gene product has been shown to bind to fungal chitin oligosaccharides and to instigate a signal cascade via the WRKYs and MAPKs pathway, leading to defence response elicitation (Zhang et al., 2007). *CERK1* gene appears to be a critical component in chitin perception as a dominant mis-sense mutation in the gene and leads to compromised non-host resistance to fungal hyphae (Lipka et al., 2007). *CERK1* has been implicated to be involved in the perception of bacterial PAMPs due to evidence of binding to peptidoglycan, which is a critical component in bacterial cell walls (Zhang et al., 2007; Buist et al., 2008; Gimenez-Ibanez et al., 2009). As a consequence of this apparent dual binding function, it appears likely that LysM gene receptors target a conserved saccharide structural domain such as GlcNAc that is present across the structurally similar PAMP candidates: lipopolysaccharides, lipochitin-oligosaccharide, peptidoglycan, β -glucans

and chitin oligosaccharides (Buist et al., 2008). These compounds are commonly found in fungi, bacteria and even viruses, and are therefore an ideal target for a PAMP receptor. As with nodule formation, the apparent broad binding range of LysM gene receptors may work in conjunction with other LysM gene products to confer pathogen binding specificity to the interaction with PAMPs to elicit an appropriate defence response.

This basal perception mechanism may have been sufficiently successful that it led to the evolution of pathogenic basal suppression, as bacterial AvrPtoB appears to target the CERK1 receptor kinase and promote bacterial virulence (Zipfel and Rathjen, 2008; Gimenez-Ibanez et al., 2009). Lux bacterial growth data in knockouts of other LysM genes (e.g. AT1G51940 and AT1G21880) (Fig. 5.9) shows similar increases in susceptibility after knockout of *CERK1*. Other suppression pathways targeting LysM genes may exist but have not yet been identified, however pathogenic selection pressure to evolve such suppression may be similar if they are as effective in conferring resistance as *CERK1*. Expansion on this hypothesised model of interaction is described in section 6.2.

Chapter 6: Discussion

6.1 Research summary

Basal resistance occurs in genetically compatible, as well as incompatible, plant-microbe interactions and acts to slow disease progression, potentially allowing completion of the host life cycle (Glazebrook et al., 2006). Basal resistance utilises receptors which detect non-race specific pathogen elicitors, known as PAMPs, and while not commonly resulting in a HR, it may share some downstream signalling and effector components with *R* gene-mediated resistance (Zipfel et al., 2004). Basal resistance mechanisms are likely polygenic and may offer more effective, durable resistance (Zipfel et al., 2004). Relatively little is known about basal disease resistance pathways and mechanisms of plant-pathogen perception which would be an important intermediate step to targeted crop improvement.

The assay using photon emission by the Lux-tagged *Pseudomonas* was shown to be directly proportional to the number of viable bacteria *in planta*. To assess the natural variation in basal disease resistance, a survey of bacterial growth of a universally virulent strain (*Pseudomonas syringae* pv. *tomato* DC3000) across 92 *Arabidopsis thaliana* ecotypes showed a substantial range of differences (~50-fold), indicating substantial natural genetic variation in basal resistance. To further investigate the genetic basis of this variation, the luminescence based protocol for high throughput quantitative assay of bacterial growth *in planta* was refined. A series of characterisation experiments, in which components of the procedure were modified around established parameters, aimed to refine the protocol and reduce data variability. When combined, these refinements and additional control comparisons reduced Lux assay measurement variation around the mean so that the variance ratios between genotypes were similar (F pr. >0.05). In addition, similarity of data points within a genotype were shown (P= >0.05) and distinctiveness between RIL genotypes was proved (P= <0.05) for the refined assay. The refined procedure was then used to test several populations of RILs to produce tightly defined bacterial growth data for accurate QTL analysis.

The original trait screen of parental ecotypes was used to select 5 RIL lines which represented a proportion of the range of observed variation (Fig.3.5). *Pst* infiltration assay screens of these lines showed a similar continuous range of variation

with no distinct data steps and variation above and below the parental means, which indicated a polygenic trait (Fig. 4.2-4.6). Initially, the Sorbo x Gy-0 RIL population was analysed for QTLs and revealed that loci around 60cM, 20cM and 40cM on chromosomes 1, 3 and 5 respectively, have major effects on the variation of basal disease resistance (Fig. 4.7). Collectively, they explain around 80% of the variation. Subsequent analysis on the Nok-3 x Ga-0 RILs revealed significant QTL peaks on chromosomes 1, 2, 3 and 4 (Fig. 4.8). A significant QTL was located on chromosome 3 in the Cvi-0 x Ag-0 RILs and two tightly defined QTL peaks were found on chromosomes 3 and 5 in the Col-0 x Ler RILs (Fig. 4.9, 4.11). In addition, significant QTL peaks were found on chromosomes 2 and 5 for the Wt-5 x Ct-1 RIL population (Fig.4.10)

The independent QTL analyses of the Sorbo x Gy-0, Nok-3 x Ga0, Cvi-0 x Ag-0 and Col-0 x Ler RIL populations showed 12 significant QTLs. Three QTL regions containing 6 significant peaks showed substantial overlap that persisted across populations on chromosomes 2 and 3 (Fig. 4.12). The RIL populations Nok-3 x Ga0 and Wt5 x Ct-1 showed similar QTL locations on chromosome 2, between 40-50 cM. Cvi-0 x Ag-0 and Col-0 x Ler and implicated the same region on chromosome 3 at 0 cM to 10 cM. Sorbo x Gy-0 and Nok-3 x Ga-0 both had significant peaks at around 20cM on chromosome 3 (Fig. 4.12).

The chromosomal regions determining the variation of basal disease resistance were delineated by markers from either side of these significant QTL peaks. Lists of the genes within these areas were then compiled and cross referenced against a list generated of all known disease and PAMP related genes based on sequence annotation and literature reports. These refined candidate gene lists significantly reduced the number of initial candidates in the region by approximately 95% (Table 4.7).

Candidate genes from these regions were categorised and cross-referenced against each of the others to determine overlap and were also assessed for their implication in biotic stresses (Fig. 4.13). The distribution frequency of individual candidates and gene families in QTL regions was assessed to determine if it was significantly higher than the background (Table 4.8). Statistical tests revealed that members from a family of 15 genes encoding receptor-like proteins with a conserved lysine motif (LysM) occurred at significantly higher rates in the QTL regions. Nine of the 12 significant QTLs discovered in the five RIL populations are closely

associated with members of the LysM family and some of them are significantly up-regulated in response to treatment with flg22 peptide (Fig. 4.15). This is consistent with the hypothesis that LysM gene products are involved in PAMP perception, leading to basal disease resistance.

QTL analysis for two populations was also repeated with a different bacterial strain (*Psm*) to indicate the specificity of the interaction. QTL graph profiles for the Nok-3 x Ga-0 RIL population show substantial similarity between *Pst* and *Psm* infections which may indicate shared perception mechanisms (Fig. 4.17). However, the *Psm* QTL profile for the Col-0 x Ler population is different to the graphs associated with the *Pst* interaction and yet, a strong correlation between LysM gene locations and QTL peaks is maintained which may indicate that combination of LysM genes may be important in conferring specificity (Fig. 4.16).

The nucleotide sequences of all the Col-0 *Arabidopsis* LysM genes were obtained from the TAIR database and used for BLAST searches to find homologous genes in other organisms as a possible indication of function. The sequences were then compiled to make phylogenetic trees to visually explain and quantify the genetic similarity between genes and close homologues (Fig. 5.3). Sequence comparison indicated a smaller subset of LysM genes with RLK function and close homology to *NFR1* and were correlated with proximity to QTL peaks. However, this is not a conclusive trend as some LysM genes that show strong affinity with QTL location are relatively poorly related with *NFR1*.

In a collaborative effort with the Volker Lab (Sainsbury Lab, UK), the LysM gene *CERK1* was sequenced from multiple genotypes to determine if any specific polymorphism was associated with differences between genotype trait data. The sequencing showed some major polymorphisms in the deduced amino acid sequence between genotypes which may change the conformational structure and possibly influence protein function. While a direct correlation between any amino acid exchanges and trait data could not be found there are multiple QTLs acting within each population which may mask a mutation's effect. Sequence data was also obtained for each of the LysM genes from both Col-0 and Ler and this variation was compared with basal trait differences, however a clear relationship between a polymorphism and trait variation could not be conclusively shown.

Use of Genevestigator meta-profiles, that combine multiple Affymetrix GeneChip *Arabidopsis* ATH1 signals arrays to give an average expression value,

showed no clear anatomical or developmental pattern to expression of these genes (Fig. 5.4, 5.5). This result was expected as receptor genes are, in general, expressed at low levels consistently throughout the plant and over its life cycle. When exposed to different biotic and abiotic stimuli, LysM expression levels showed a number of substantial alterations, such as similar expression rates of LysM genes when exposed to strains of *P.syringae* and chitin (previously indicated as a LysM binding factor) and down-regulation during certain abiotic stress treatments (Ball et al., 2004) (Fig. 5.6). These results are consistent with the hypothesis of LysM involvement as an upstream receptor or signalling component, although such data is not conclusive.

A *Pseudomonas* strain carrying *AvrPtoB* has been described as having a kinase suppression function influencing a tomato orthologue of the *CERK1* LysM gene (Miya et al., 2007). It is important to determine if a correlation could be shown between QTLs, LysM gene location and suppression of basal resistance. *AvrPtoB* deletion mutant strain was shown to have moderately (not significantly ($P = >0.05$) for most lines) decreased growth rates compared with wild-type *Pst* (Fig. 4.18). Examination using different ecotypes and analysis of QTL locations indicated that *AvrPtoB* kinase suppression had some influence on basal resistance. This may affect several LysM genes and may not be limited to the *CERK1* gene as previously reported (Gimenez-Ibanez et al., 2009).

A double-blind experiment utilising *Pst* infiltration bacterial growth tests on T-DNA knockout lines with insertions in individual LysM RLKs aimed to establish a direct effect on the basal resistance. Two paired PCR amplification reactions were performed to confirm the insertion genotype (WT, heterozygous, homozygous) as some of the lines were still segregating (Fig. 5.8). Homozygous knockout lines show a moderate increase in susceptibility (40-60%) for lines targeting specific LysM genes, compared with Col-0 controls (Fig. 5.9). Not all lines show an increased susceptibility pattern but most tested lines show a significant effect on basal trait after gene knockout. LysM genes are therefore directly implicated in basal resistance. Further assessment to confirm a reduction in LysM gene transcription levels would be needed to draw a definitive conclusion.

6.2 LysM gene family and PAMP recognition

The search for PRR and their associated PAMP molecules has intensified in recent literature and the *CERK1* LysM RLK gene has been recently implicated in perception of pathogenic fungi (Miya et al., 2007). Chitin oligosaccharides derived from fungal hyphae reportedly interact with *CERK1*, resulting in a signalling cascade via the WRKYs and MAPKs pathway (Wan et al., 2008). A dominant mis-sense mutation of the *CERK1* gene shows fully compromised non-host resistance to fungal hyphae and is independent of *BAK1* and *BKK1*, which distinguishes this pathway from the *FLS2* upstream mechanism (Wan et al., 2008; Gimenez-Ibanez et al., 2009). Downstream utilisation of similar signalling with the *FLS2* and *R* gene-mediated responses is likely to be due to conservation MAPK mediated transduction (Kaku et al., 2006).

Prior to literature associated with chitin perception, *Arabidopsis* LysM genes had no described function and were commonly mentioned in association with being orthologous with Nod factor recognition genes *NFR1* and *NFR5* from *Lotus japonicus* (Radutoiu et al., 2003). Nod factors are extra-cellular lipo-chito-oligosaccharides (LCOs) induced following exposure to plant flavanoids in the rhizosphere and consist of a chitin oligomeric backbone, which typically comprises of 3 to 5 N-acetylglucosamine (GlcNAc) molecules (D'Haeze and Holsters, 2002). Functional group substitutions can occur at terminal or non-terminal residues with an N-acylated group attached to the non-reducing terminal saccharide residue (D'Haeze and Holsters, 2002). *NFR1* and *NFR5* LysM genes are essential to host-symbiont specificity and are involved with recognition of specific Nod factor chemical structure to detect bacterial presence and appropriately initiate nodulation if suitable rhizobacteria are perceived (Gough, 2003; Radutoiu et al., 2003). Like the *Arabidopsis* LysM genes, *NFR1* and *NFR5* are also membrane-bound, have receptor-like kinase function and contain serine/ threonine receptors that work in conjunction to determine Nod factor specificity in roots (Gough, 2003; Madsen et al., 2003). *NFR1* and *NFR5* specifically recognise rhizobial bacteria using multiple LysM domains with the LysM2 domain from *NFR5*, largely responsible for specificity (Radutoiu et al., 2007). A single amino-acid substitution in the LysM2 domain results in drastically altered nodulation efficiencies (Radutoiu et al., 2007). Modelling predicts a cleft for Nod-factor binding at this domain with the L/K118 amino acid, shown *in vivo* to aid recognition of Nod-factor structure, positioned in close proximity to the groove

(Radutoiu et al., 2007). Similar reductions are observed during transcriptional imbalance between *NFR1* and *NFR5* which adds to the hypothesis that these gene products are part of the same receptor complex (Radutoiu et al., 2007). The working hypothesis of this project predicts a similar evolutionary relationship between these pathways in PAMP perception specificity at the cell surface. Downstream signal alterations, which in the legume symbiosis results in nodule formation, would be modified to activate defence reactions for basal resistance.

Via knockout analysis of the Rice LysM gene *CEBiP*, it was shown that an orthologue to *CERK1* LysM Arabidopsis gene is an important component in the activation of elicitor induced oxidative burst as well as perception and transduction of chitin oligosaccharide (Kaku et al., 2006). *CEBiP* is thought to function as a cell surface receptor for chitin elicitor and RNAi silencing of this gene shows a reduction in chitin binding (Zipfel, 2009). This similarity suggests that defence pathways based on perception of chitin through LysM genes are widely conserved amongst plant species (Kaku et al., 2006). The LysM perception component therefore represents a conserved link between symbiosis and defence pathways. If this similarity is further conserved then, as hypothesised, the LysM genes may act in conjunction with other LysM genes to confer specificity (as with *NFR1* and *NFR5* LysM genes) to the PAMP target that LysM gene based receptors may interact with (Knogge and Scheel, 2006). The details of any interactions are not yet known, although yeast-two-hybrid experiments suggest formation of a hetero-dimer between the two gene products (Kaku et al., 2006). This is mechanistically feasible as transmembrane *CEBiP* protein lacks intracellular signalling domains, so may require an interaction with *CERK1* orthologues or other LysM genes to transduce the chitin signal (Schwessinger and Zipfel, 2008). Earlier QTL analysis peak locations may be a useful indication of which LysM genes are working in conjunction in *Arabidopsis* genotypes for a specific pathogen.

Following perception of Nod-factor molecules in wild-type root nodulation formation, root hair curling and deformation is easily observable within 24hrs of inoculation. Double mutants of LysM-type serine/threonine receptor kinase genes in Lotus, *nfr1* and *nfr5*, show no cellular alterations when exposed to *M.loti* bacteria or purified lipochitin-oligosaccharide which suggests *NFR1* and *NFR5* genes are involved early in the Nod-factor perception interaction (Radutoiu et al., 2007). The double and single *nfr1* and *nfr5* mutant cross-comparison shows that both the *NFR1*

and *NFR5* genes are essential for mounting the earliest responses to Nod-factor via the SYMRK kinase complex (Radutoiu *et al.*, 2003). Double mutant analysis of LysM genes in *Arabidopsis* could be constructed by crossing two mutants homozygous for different mutations to test the hypothesis that LysM genes maybe working in conjunction in a similar mechanistic way to *NFR1* and *NFR5*. As described in section 6.5, this was attempted but was unsuccessful in producing progeny with double LysM gene knockout.

The exact function of LysM domains is not clear however they frequently occur in bacteriophage proteins, peptidoglycan hydrolases, peptidases and chitinases which are all compounds which require specific differentiation between pathogen and host (Zhang *et al.*, 2007). Currently, proteins containing one or more LysM domains have been found in 4000 eukaryotes and prokaryotes however none have been observed in archaeobacteria, which presumably deviated from the main prokaryote phyla prior to the evolution of LysM domains (Zhang *et al.*, 2007). Therefore, LysM domains are highly conserved sequences separated by flexible lengths of spacing sequences of relatively low homology usually containing serine, threonine, aspartic acid and proline, which presumably act to position the LysM binding domains to the target substrate (Buist *et al.*, 2008).

Peptidoglycan, derived from bacteria, constitutes a PAMP which initiates an immune response in *Arabidopsis* resulting in: elevation of reactive oxygen and nitrogen species, increases in cytoplasmic calcium, increased callose and induction of MAPK activities (Gust *et al.*, 2007). The minimal chemical constituent of peptidoglycan structure found to trigger a PAMP response in animals is muramyl dipeptide (MDP) (Gust *et al.*, 2007). Transmembrane LRR-RLK proteins are a common cellular component and share striking structural similarity in both plant and animal perception mechanisms which may suggest a common evolutionary past and possible similarity in current pathogen perception mechanisms. For example, *FLS2* resembles the extracytoplasmic domain of human TLR5 reportedly capable of bacterial flagellin recognition and *Arabidopsis EFR* has mammalian peptidoglycan receptor homologues; *NOD1* and *NOD2* (Hayashi *et al.*, 2001; Inohara and Nunez, 2003). In addition to the evidence linking LysM genes to fungal chitin recognition, recent data from binding assays show Gram-positive bacteria implicated in LysM non-covalent binding to peptidoglycan molecules and structurally similar viral glycoprotein (Buist *et al.*, 2008).

Evidence from a tomato orthologue of *CERK1* also links LysM genes with bacterial perception as the gene product appears to be a potential target for suppression by the AvrPtoB protein secreted by *Pseudomonas* and this was conclusively shown in a recent paper (Gimenez-Ibanez et al., 2009). *CERK1* is shown to be a determinant of bacterial immunity, however bacterial AvrPtoB ubiquitinates the kinase domain *in vitro* and is shown to initiate degradation *in vivo* which blocks a defence response via this receptor (Gimenez-Ibanez et al., 2009). This project aimed to extend these experiments by measuring susceptibility with Lux trials which included different *Arabidopsis* ecotypes. In comparison with *Pst* Δ AvrPtoB, these trials show significant increases in susceptibility for the Nok-3 ecotype exposed to *Pst* DC3000 but a small ($P = >0.05$) increase in bacteria for Col-0 and other ecotypes (Fig. 4.18). Ecotype variations in AvrPtoB effect were also observed between Col-0 and Ws-4 by (Gimenez-Ibanez et al., 2009). The results also suggest that in comparison with QTL peak locations, suppression across ecotypes is wide spread but may not be limited to *CERK1*, and AvrPtoB may suppress kinase function in multiple LysM genes.

To explain the duality of recognition between fungal and bacterial sources involving LysM genes, it was previously hypothesised that the PAMP recognition target for LysM genes may be one of several structurally similar compounds: lipopolysaccharides (LPS), lipochitin-oligosaccharide, peptidoglycan, β -glucans and chitin oligosaccharides (Buist et al., 2008). These LysM PAMP binding candidates are all extracellular polymeric substances (EPS) which are components of the protective bacterial biofilm layer or components of the cell wall (Plude et al., 1991). LysM gene products have also been shown to have some binding interaction with viral glycoprotein which would also add support to the hypothesis that LysM genes are acting as a broad ranging receptor for PAMPs (Buist et al., 2008).

LysM genes have been shown to interact with cellular components from a diverse array of species. Preliminary structural comparison suggests it is possible that LysM genes may be targeting a conserved structural domain or amino acid sequence that links at least two of these candidates (Buist et al., 2008). A potential conserved domain would be the amide, N-acetylglucosamine (GlcNAc), which is a monomeric unit which can link to form chitin polymer chains or can cross-link with alternating units of N-acetylmuramic acid to create a lattice structure called peptidoglycan (Hedrick et al., 1988). If the GlcNAc or similar conserved sub-units of bacterial cell

wall structures were the binding target of LysM genes, then this would effectively link and explain the diverse binding and species target evidence emerging (Miya et al., 2007; Buist et al., 2008).

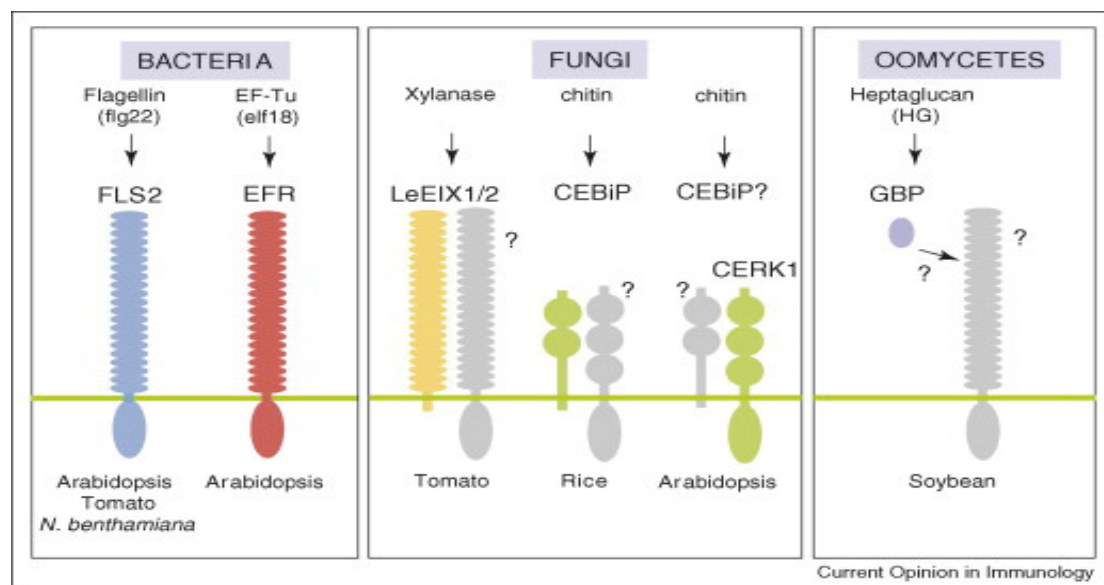


Figure 6.1 A simplified summary of the interactions between *FLS2*, *BAK1* and other components in PAMP recognition. Figure shows the formation of a complex between *FLS2* and *BAK1* that result in signal transduction and basal defence. *BAK1* and *BKK1* have also been implicated as a regulatory facilitator in *EFR* PAMP perception and regulation of HR response which is typically suppressed during a basal response. Figure published by (Zipfel, 2008) with permission of Elsevier publishing.

6.3 Progression of the project and subject area

Due to the highly dynamic nature and rapid expansion of scientific knowledge, it is useful to place this project into context within which it was conducted. At the start of this project in October 2005, basal resistance had been identified as an underdeveloped area of interest with the potential to offer durable disease resistance, with two PAMP systems having been clearly identified; EF-Tu/-EFR and flg22/-FLS2. The LRR-RLK EFR PAMP receptor in *Arabidopsis* perceives the abundant bacterial elongation factor EF-Tu epitope elf18, leading to a signalling cascade via MAPK and WRKY genes which activates basal defence responses (section 1.3.1) (Kunze et al., 2004). FLS2 is a membrane bound LRR-RLK gene involved in the perception of the flagellin peptide flg22 (Zipfel *et al.*, 2004). The FLS2 receptor forms a complex with BAK1 to initiate basal plant defence response via the MAP

kinase pathway (section 1.3.2) (Chinchilla et al., 2007). The distinction between resistance mechanisms is still unclear, although it was considered highly likely that even with these recent PAMP system examples, many components of this basal defence were still unknown.

The LysM gene family was identified as an area of possible interest for this study in 2006 using preliminary QTL data, however, at that time there were no publications which described a function for LysM genes in basal resistance. At this point, literature search engines only produced references to LysM gene homology to *NFR1* in *Lotus japonicus*, and TAIR website annotation information listed all LysM genes as having unknown function. During 2006 and 2007, the Lux assay was refined to improve accuracy and multiple RIL line crosses were tested and independent QTL graphs generated which coincide with the locations of LysM genes. During 2008 and 2009, the LysM family was being assessed in this project for its potential to be involved in bacterial PAMP perception using sequence and phylogenetic analysis, microarray based expression profiling and T-DNA knockouts of LysM genes to determine effect on bacterial growth *in planta*. The project officially ended in September 2009. The publication record has changed considerably during this project with the implication of the LysM gene *CERK1* being involved in the perception of fungal chitin (Kaku et al., 2006; Wan et al., 2008). More recently, LysM domains were linked with peptidoglycan binding, a major component of bacterial cell walls (Buist et al., 2008). Current literature frequently describe the *CERK1* gene and its fungal chitin interactions, however, this project presents evidence for the hypothesis that many members in the family, including *CERK1*, also have a role in bacterial PAMP perception. To place the project's findings into context, and to update the current hypothesis of action, recent findings relating to the LysM gene family are summarised below.

6.4 Application of the findings and commercialisation potential

Current understanding of the mechanistic and genetic components involved in controlling basal resistance is not sufficient to selectively improve resistance for commercial application. Such understanding is a critical prerequisite to the enhancement of basal resistance and its application to commercial varieties. This

project aimed to improve the understanding of the genetic basis of basal resistance and, through QTL analysis and knockout lines, has implicated the LysM gene family to be involved. This study is at an early stage of development and further validation is required to unequivocally link LysM genes as influential players in the PAMP/ PTI perception. The CASE sponsor of this study, Plant Bioscience Limited (PBL), does have a current patent relating to the *CERKI* gene. It may be prudent for them to consider expanding the patent to a wider range of LysM genes in the future after further study to define their mechanistic function. The original patent linked *CERKI* specifically to fungal perception and so would be limited to this application whereas current literature and knockout experiments in this project suggest a broader perception mechanism extending to *Pseudomonas* bacteria. The implication of LysM gene family involvement in basal resistance was derived independently, prior to the knowledge of the PBL *CERKI* patent.

6.5 Future work

Our appreciation and understanding of this area of science has recently expanded rapidly. However, there is still a huge amount that is unclear or completely unknown. Continuation of this project does have a high probability of yielding interesting results and advancing our knowledge, as there are several areas which need to be examined more conclusively. This project and some published evidence suggest that LysM genes, particularly those with RLK function, may have a role in basal resistance. Future work needs to provide evidence to unequivocally validated LysM genes as a player in PAMP mediated resistance and this will require evidence from multiple experiments. Such evidence may be obtained using a variety and combination of experimental approaches, some of which are described below, to elucidate the mechanisms and role of LysM genes in the PTI pathway.

RT-qPCR has a potentially wide range of applications and has previously been used to assess gene transcript levels and to validate microarray experiment readings, in addition to assessing whether transcript levels had been significantly disrupted in knockout lines (Ginzinger, 2002). Quantitative PCR was used in the elucidation of the DELLA proteins, such as RGL1, RGL2, and RGL3, typically associated with seed germination and flower development (Tyler et al., 2004). Knockout lines containing

T-DNA inserts were extensively used to characterise the effect an individual gene has on final trait function and RT-qPCR was used to demonstrate that transcript levels were reduced in the majority of T-DNA lines (Tyler et al., 2004). However, although RT-qPCR is a sensitive technique, if used to examine transcript levels of LysM genes, the procedures would need multiple calibration controls as the differences are expected to be low based on the transcriptional profiles described in section 5.4. The quantitative assessment of transcript level may also be used for pathway dissection as was used following the observation that *SEN1*, a gene typically up-regulated prior to senescence, became constitutively expressed in a *WRKY6* knockout mutant. That would suggest that in the wild-type plant, the WRKY6 protein might act as a negative regulator of *SEN1* (Hanaoka et al., 2002; Schenk et al., 2005). Assessment of *LysM* gene transcription rates in knockout lines may help to order the hierarchy of events in the perception pathway or influential factors regulating expression or effect.

The technique of using single T-DNA knockouts to disrupt gene transcription can be continued to assess multi-gene traits through multiple knockout mutations (Krysan et al., 1999). Additional crossing of homozygous knockout lines to pyramid the LysM RLK knockouts can be confirmed with PCR and followed by *Pst* infiltration assay to establish any additional effect on the trait. A similar approach in *Arabidopsis* has been used to assess the function of amylase-like proteins AtAMY1, AtAMY2, and AtAMY3, suspected to be involved in starch breakdown. A triple T-DNA knock-out line was generated and shown to contain degraded starch in the leaf comparable to wild-type, therefore these amylase-like proteins were concluded to be unnecessary for transitory starch breakdown (Yu et al., 2005). An attempt was made to cross homozygous T-DNA knockout lines to ‘pyramid’ the LysM RLK knockouts and test the effect following *Pst* infiltration. Parental knockout lines showed significant differences between insertion line bacterial growth and controls were selected from different insertions in LysM RLKs. Seeds were generated by crossing the homozygous *Arabidopsis* T-DNA insertion lines. The resulting progeny were assessed for double insertion, using two sets of the paired reactions used previously, to separately assess both T-DNA insertions within each of the target genes. Initial seed production from the crosses was poor, as were the germination rates of the harvested seed. However, the generated seeds and resulting plants were Lux *Pst* infiltrated to test basal resistance. Unfortunately, none of the plants or lines contained a double knockout and time constraints prevented additional attempts. T-DNA

mutagenesis and particularly multiple T-DNA insertion has previously produced some potentially detrimental side effects such as minor base substitutions, additions and deletions or rare chromosomal translocations (Tax and Vernon, 2001). Despite the initial difficulties with producing a successful double knockout mutant in this project, the technique of crossing homozygous single knockout mutants has proved successful in establishing function by disrupting multiple GID1 receptor genes with the result of completely shutting down the gibberellin signal in *Arabidopsis* (Iuchi et al., 2007).

An additional use of T-DNA insertions would be LysM gene candidate complementation experiments, which attempt to enhance or restore function in susceptible lines through the introduction and over-expression of specific candidate genes. This technique was used to test the function of the *WRKY18* gene (Chen and Chen, 2002). Using moderate transcription promoters, the plants exhibited an increase in *PR* gene expression and enhanced resistance to *Pseudomonas syringae* compared to wild-type. In association with *NPR1/NIM1*, the *WRKY18* gene product was shown to positively modulate defence-related gene expression and disease resistance (Chen and Chen, 2002). Increase of the transcription rates may not necessarily correlate with enhancement of the trait and when constitutively over-expressed with the 35S promoter, it may result in deleterious physiological effects such as stunted growth, as occurred in this example (Chen and Chen, 2002). Therefore, over-expression may give a useful indication of gene function but if expression rates are not within the range of gene expression rates typically observed within the population, then any effect is difficult to attribute to gene action.

Compelling evidence could be obtained for LysM gene family involvement if the study went back to the QTL data and RIL lines in an effort to Mendelise the trait. This approach would make the assumption that resistance mediated by a LysM gene follows simple Mendelian segregation and the addition of a single LysM gene locus can confer enhanced resistance. In RIL populations, which indicate only 1 QTL, crosses could be conducted to reintroduce the implicated chromosomal section to establish if this restored trait phenotype scores. To generate crosses accounting for multiple QTL peaks would be unfeasibly time consuming for this project. Also, PAMP resistance is likely controlled by multiple loci which may undermine the stark phenotypic differences needed for Mendelising a trait through crosses. A similar approach has been used for *R-gene* defence traits, such as resistance to Barley stripe rust, where QTL pyramiding has provided a link between trait function and specific

gene control of the trait as pyramiding of major QTLs into susceptible lines restores some, or all, of the resistance (Castro et al., 2003b; Castro et al., 2003a). Marker mapping of Arabidopsis lines can facilitate the crossing of significant LysM QTL peak regions into RI lines which do not indicate those regions to establish if this has an effect on the trait. This evidence could be supported by the cloning and transfer LysM genes into susceptible / resistance genotypes to confer enhanced resistance and further link QTL peaks with individual genes' effects on trait variation.

Genetic dissection of trait pathways has previously been achieved using microarray analysis of global expression to identify and characterise the function components of floral induction such as *SPL* and *LFY* genes (Schmid et al., 2003b). Microarray analysis has helped elucidate that the salicylate and jasmonate pathways, previously thought to act antagonistically, are involved in a more complex network of regulatory interactions and coordination between different defence signalling pathways (Schenk et al., 2000). The expression of over two thousand genes was examined following exposure to an incompatible fungal pathogen, *Alternaria brassicicola*, or exposure to plant hormones SA, MeJA and ET. Significant changes in regulation were found in 106 genes with no previously described function or homology to known defence genes which could then be assessed for direct pathway involvement (Schenk et al., 2000). Additional components of the basal resistance pathway, which may have been overlooked or discounted by the refinement methods used in this project, could be identified using microarrays (Ossowski et al., 2008). Comparison between mock vs. infiltrated tissue at different time points may identify gene candidates which are highly transcriptionally activated in their response to bacterial infection. LysM genes do not typically show high transcriptional response following exposure to flg22 however, other components in any LysM mediated basal resistance pathway may show greater levels of transcriptional activation, and so would indicate potential components in the pathway (Zipfel et al., 2004).

The *CERK1* LysM gene reportedly binds to fungal chitin however, it is likely that bacterial AvrPtoB protein targets this gene for suppression and other evidence suggests that LysM binds to bacterial peptidoglycan (Miya et al., 2007; Buist et al., 2008; Gimenez-Ibanez et al., 2009). There remains a discrepancy between likely PAMP perception targets, as evidence from this project and others also suggests *CERK1* is involved in bacterial perception as well as from fungal chitin. A conserved structural motif may exist between potential candidate PAMPs which are structurally

similar, such as: lipopolysaccharides, lipochitin-oligosaccharide, peptidoglycan, B-glucans and chitin oligosaccharides (Buist et al., 2008). Previous studies in animals have performed similar binding assays by purifying peptidoglycan and showing affinity to the PGRP protein using the cascade of prophenoloxidase as an assay (Yoshida et al., 1996). An alternative approach used paralogue LysM domain *N*-acetylglucosaminidase (AcmA) fusion protein, which binds specifically to the C-terminal LysM domain, combined with *L. lactis* cells to identify purified peptidoglycan as the component to which the LysM domain of AcmA binds (Buist et al., 2008). Binding assays may be helpful in determining whether peptidoglycan, chitin or any of the previously suggested substrates also bind to the other LysM genes which have been implicated by this project to be involved in basal resistance.

References

- Aarts, N., Metz, M., Holub, E., Staskawicz, B.J., Daniels, M.J., and Parker, J.E. (1998). Different requirements for EDS1 and NDR1 by disease resistance genes define at least two R gene-mediated signaling pathways in Arabidopsis. *Proc Natl Acad Sci USA* **95**, 10306-10311.
- Abramovitch, R.B., Anderson, J.C., and Martin, G.B. (2006a). Bacterial elicitation and evasion of plant innate immunity. *Nat Rev Mol Cell Biol* **7**, 601-611.
- Abramovitch, R.B., Janjusevic, R., Stebbins, C.E., and Martin, G.B. (2006b). Type III effector AvrPtoB requires intrinsic E3 ubiquitin ligase activity to suppress plant cell death and immunity. *Proc Natl Acad Sci U S A* **103**, 2851-2856.
- Abramovitch, R.B., Kim, Y.J., Chen, S., Dickman, M.B., and Martin, G.B. (2003). Pseudomonas type III effector AvrPtoB induces plant disease susceptibility by inhibition of host programmed cell death. *Embo J* **22**, 60-69.
- Alonso-Blanco, C., and Koornneef, M. (2000). Naturally occurring variation in Arabidopsis: an underexploited resource for plant genetics. *Trends Plant Sci* **5**, 22-29.
- Arends, D., Prins, P., Broman, K., and Jansen, R. (2010). Tutorial - Multiple-QTL Mapping (MQM) Analysis. *RQTL*.
- Arrighi, J.F., Barre, A., Ben Amor, B., Bersoult, A., Soriano, L.C., Mirabella, R., de Carvalho-Niebel, F., Journet, E.P., Gherardi, M., Huguet, T., Geurts, R., Denarie, J., Rouge, P., and Gough, C. (2006). The Medicago truncatula lysin [corrected] motif-receptor-like kinase gene family includes NFP and new nodule-expressed genes. *Plant Physiol* **142**, 265-279.
- Asins, M.J. (2002). Present and future of quantitative trait locus analysis in plant breeding. *Plant Breeding* **121**, 281-291.
- Asselbergh, B., and Hofte, M. (2007). Basal tomato defences to Botrytis cinerea include abscisic acid-dependent callose formation. *Physiol Mol Plant P* **71**, 33-40.
- Baker, N.R. (2008). Chlorophyll fluorescence: a probe of photosynthesis in vivo. *Annu Rev Plant Biol* **59**, 89-113.
- Baldauf, S.L. (2003). Phylogeny for the faint of heart: a tutorial. *Trends Genet* **19**, 345-351.
- Ball, L., Accotto, G.P., Bechtold, U., Creissen, G., Funck, D., Jimenez, A., Kular, B., Leyland, N., Mejia-Carranza, J., Reynolds, H., Karpinski, S., and Mullineaux, P.M. (2004). Evidence for a direct link between glutathione biosynthesis and stress defense gene expression in Arabidopsis. *Plant Cell* **16**, 2448-2462.
- Bateman, A., and Bycroft, M. (2000). The structure of a LysM domain from E-coli membrane-bound lytic murein transglycosylase D (MltD). *J Mol Biol* **299**, 1113-1119.
- Bednarek, P., and Osbourn, A. (2009). Plant-microbe interactions: chemical diversity in plant defense. *Science* **324**, 746-748.
- Belkhadir, Y., Subramaniam, R., and Dangl, J.L. (2004). Plant disease resistance protein signaling: NBS-LRR proteins and their partners. *Curr Opin Plant Biol* **7**, 391-399.
- Bittel, P., and Robatzek, S. (2007). Microbe-associated molecular patterns (MAMPs) probe plant immunity. *Curr Opin Plant Biol* **10**, 335-341.

- Bouarab, K., Melton, R., Peart, J., Baulcombe, D., and Osbourn, A.** (2002). A saponin-detoxifying enzyme mediates suppression of plant defences. *Nature* **418**, 889-892.
- Boyes, D.C., Nam, J., and Dangel, J.L.** (1998). The *Arabidopsis thaliana* RPM1 disease resistance gene product is a peripheral plasma membrane protein that is degraded coincident with the hypersensitive response. *Proc Natl Acad Sci U S A* **95**, 15849-15854.
- Brisson, L.F., Tenhaken, R., and Lamb, C.** (1994). Function of Oxidative Cross-Linking of Cell Wall Structural Proteins in Plant Disease Resistance. *Plant Cell* **6**, 1703-1712.
- Buist, G., Steen, A., Kok, J., and Kuipers, O.P.** (2008). LysM, a widely distributed protein motif for binding to (peptido)glycans. *Mol Microbiol* **68**, 838-847.
- Burton, H., Dye, N., and Bush, L.** (1992). Distribution of tobacco constituents in tobacco leaf tissue. 1. Tobacco-specific nitrosamines, nitrate, nitrite, and alkaloids. *J. Agric. Food Chem* **40** (6), pp 1050–1055.
- Caicedo, A.L., Schaal, B.A., and Kunkel, B.N.** (1999). Diversity and molecular evolution of the RPS2 resistance gene in *Arabidopsis thaliana*. *P Natl Acad Sci USA* **96**, 302-306.
- Capoen, W., Den Herder, J., Rombauts, S., De Gussem, J., De Keyser, A., Holsters, M., and Goormachtig, S.** (2007). Comparative transcriptome analysis reveals common and specific tags for root hair and crack-entry invasion in *Sesbania rostrata*. *Plant Physiol* **144**, 1878-1889.
- Castro, A.J., Chen, X.M., Hayes, P.M., and Johnston, M.** (2003a). Pyramiding quantitative trait locus (QTL) alleles determining resistance to barley stripe rust: Effects on resistance at the seedling stage. *Crop Sci* **43**, 651-659.
- Castro, A.J., Chen, X.M., Corey, A., Filichkina, T., Hayes, P.M., Mundt, C., Richardson, K., Sandoval-Islas, S., and Vivar, H.** (2003b). Pyramiding and validation of quantitative trait locus (QTL) alleles determining resistance to barley stripe rust: Effects on adult plant resistance. *Crop Sci* **43**, 2234-2239.
- Chelkowski, J., and Koczyk, G.** (2003). Resistance gene analogues of *Arabidopsis thaliana*: recognition by structure. *J Appl Genet* **44**, 311-321.
- Chen, C.H., and Chen, Z.X.** (2002). Potentiation of developmentally regulated plant defense response by AtWRKY18, a pathogen-induced *Arabidopsis* transcription factor. *Plant Physiol* **129**, 706-716.
- Cheong, Y.H., Chang, H.S., Gupta, R., Wang, X., Zhu, T., and Luan, S.** (2002). Transcriptional profiling reveals novel interactions between wounding, pathogen, abiotic stress, and hormonal responses in *Arabidopsis*. *Plant Physiol* **129**, 661-677.
- Chinchilla, D., Zipfel, C., Robatzek, S., Kemmerling, B., Nurnberger, T., Jones, J.D., Felix, G., and Boller, T.** (2007). A flagellin-induced complex of the receptor FLS2 and BAK1 initiates plant defence. *Nature* **448**, 497-500.
- Chisholm, S., Coaker, G., Day, B., and Staskawicz.** (2006). Host-Microbe Interactions: Shaping the evolution of the plant immune response. *Cell* **124**, 803-814.
- Churchill, G.A., and Doerge, R.W.** (1994). Empirical Threshold Values for Quantitative Triat Mapping. *Genetics* **138**, 963-971.
- Clarke, J.D., Volko, S.M., Ledford, H., Ausubel, F.M., and Dong, X.** (2000). Roles of salicylic acid, jasmonic acid, and ethylene in cpr-induced resistance in *arabidopsis*. *Plant Cell* **12**, 2175-2190.

- Clarke, J.H., Mithen, R., Brown, J.K.M., and Dean, C.** (1995). Qtl Analysis of Flowering Time in *Arabidopsis-Thaliana*. *Mol Gen Genet* **248**, 278-286.
- Collmer, A., Badel, J.L., Charkowski, A.O., Deng, W.L., Fouts, D.E., Ramos, A.R., Rehm, A.H., Anderson, D.M., Schneewind, O., van Dijk, K., and Alfano, J.R.** (2000). *Pseudomonas syringae* Hrp type III secretion system and effector proteins. *Proc Natl Acad Sci U S A* **97**, 8770-8777.
- Craig, A., Ewan, R., Mesmar, J., Gudipati, V., and Sadanandom, A.** (2009). E3 ubiquitin ligases and plant innate immunity. *J Exp Bot* **60**, 1123-1132.
- Creelman, R.A., Tierney, M.L., and Mullet, J.E.** (1992). Jasmonic Acid Methyl Jasmonate Accumulate in Wounded Soybean Hypocotyls and Modulate Wound Gene-Expression. *P Natl Acad Sci USA* **89**, 4938-4941.
- D'Haese, W., and Holsters, M.** (2002). Nod factor structures, responses, and perception during initiation of nodule development. *Glycobiology* **12**, 79r-105r.
- Dangl, J.L., and Jones, J.D.G.** (2001). Plant pathogens and integrated defence responses to infection. *Nature* **411**, 826-833.
- Davis, K., Schoot, E., and Ausubel, F.M.** (1991). Virulence of selected phytopathogenic pseudomonads in *Arabidopsis thaliana*. *Mol Plant Microbe In* Vol. 4, No. 5, pp. 477-488.
- de Meaux, J., and Mitchell-Olds, T.** (2003). Evolution of plant resistance at the molecular level: ecological context of species interactions. *Heredity* **91**, 345-352.
- Dean, C., Sjodin, C., Bancroft, I., Lawson, E., Lister, C., Scofield, S., and Jones, J.** (1991). Development of an efficient transposon tagging system in *Arabidopsis thaliana*. *Symp Soc Exp Biol* **45**, 63-75.
- Deslandes, L., Olivier, J., Peeters, N., Feng, D.X., Khounloham, M., Boucher, C., Somssich, I., Genin, S., and Marco, Y.** (2003). Physical interaction between RRS1-R, a protein conferring resistance to bacterial wilt, and PopP2, a type III effector targeted to the plant nucleus. *Proc Natl Acad Sci U S A* **100**, 8024-8029.
- Despres, C., Chubak, C., Rochon, A., Clark, R., Bethune, T., Desveaux, D., and Fobert, P.R.** (2003). The *Arabidopsis* NPR1 disease resistance protein is a novel cofactor that confers redox regulation of DNA binding activity to the basic domain/leucine zipper transcription factor TGA1. *Plant Cell* **15**, 2181-2191.
- Devoto, A., and Turner, J.G.** (2003). Regulation of jasmonate-mediated plant responses in *Arabidopsis*. *Ann Bot-London* **92**, 329-337.
- Efron, B., Halloran, E., and Holmes, S.** (1996). Bootstrap confidence levels for phylogenetic trees. *P Natl Acad Sci USA* **93**, 7085-7090.
- Ellis, J., Dodds, P., and Pryor, T.** (2000). Structure, function and evolution of plant disease resistance genes. *Curr Opin Plant Biol* **3**, 278-284.
- Espinosa, A., and Alfano, J.R.** (2004). Disabling surveillance: bacterial type III secretion system effectors that suppress innate immunity. *Cell Microbiol* **6**, 1027-1040.
- Eulgem, T.** (2006). Dissecting the WRKY web of plant defense regulators. *PLoS Pathog* **2**, e126.
- Falconer, D.S., and Mackay, T.F.C.** (1996). *Introduction to Quantitative Genetics*. (Prentice Hall).
- Fan, J., Crooks, C., and Lamb, C.** (2008). High-throughput quantitative luminescence assay of the growth in planta of *Pseudomonas syringae*

- chromosomally tagged with *Photobacterium luminescens* luxCDABE. *Plant J* **53**, 393-399.
- Fan, J., Hill, L., Crooks, C., Doerner, P., and Lamb, C.** (2009). Absciscic acid has a key role in modulating diverse plant-pathogen interactions. *Plant Physiol* **150**, 1750-1761.
- Fell, D.A., and Thomas, S.** (1995). Physiological Control of Metabolic Flux - the Requirement for Multisite Modulation. *Biochem J* **311**, 35-39.
- Feys, B.J., and Parker, J.E.** (2000). Interplay of signaling pathways in plant disease resistance. *Trends Genet* **16**, 449-455.
- Feys, B.J., Moisan, L.J., Newman, M.A., and Parker, J.E.** (2001). Direct interaction between the Arabidopsis disease resistance signaling proteins, EDS1 and PAD4. *Embo Journal* **20**, 5400-5411.
- Flor, H.H.** (1955). Host-Parasite Interaction in Flax Rust - Its Genetics and Other Implications. *Phytopathology* **45**, 680-685.
- Flors, V., Ton, J., van Doorn, R., Jakab, G., Garcia-Agustin, P., and Mauch-Mani, B.** (2008). Interplay between JA, SA and ABA signalling during basal and induced resistance against *Pseudomonas syringae* and *Alternaria brassicicola*. *Plant J* **54**, 81-92.
- Fouts, D.E., Badel, J.L., Ramos, A.R., Rapp, R.A., and Collmer, A.** (2003). A *pseudomonas syringae* pv. tomato DC3000 Hrp (Type III secretion) deletion mutant expressing the Hrp system of bean pathogen *P. syringae* pv. *syringae* 61 retains normal host specificity for tomato. *Mol Plant Microbe Interact* **16**, 43-52.
- Gale, M.D., and Devos, K.M.** (1998). Plant comparative genetics after 10 years. *Science* **282**, 656-659.
- Gassmann, W.** (2005). Natural variation in the Arabidopsis response to the avirulence gene *hopPsyA* uncouples the hypersensitive response from disease resistance. *Mol Plant Microbe Interact* **18**, 1054-1060.
- Gazzani, S., Gendall, A.R., Lister, C., and Dean, C.** (2003). Analysis of the molecular basis of flowering time variation in Arabidopsis accessions. *Plant Physiol* **132**, 1107-1114.
- Gimenez-Ibanez, S., Hann, D.R., Ntoukakls, V., Petutschnig, E., Lipka, V., and Rathjen, J.P.** (2009). *AvrPtoB* Targets the LysM Receptor Kinase CERK1 to Promote Bacterial Virulence on Plants. *Current Biology* **19**, 423-429.
- Ginzinger, D.** (2002). Gene quantification using real-time quantitative PCR: An emerging technology hits the mainstream. *Experimental Hematology* **30** (2002) 503-512.
- Glazebrook, J.** (2001). Genes controlling expression of defense responses in Arabidopsis - 2001 status. *Curr Opin Plant Biol* **4**, 301-308.
- Glazebrook, J., Rogers, E.E., and Ausubel, F.M.** (1997). Use of Arabidopsis for genetic dissection of plant defense responses. *Annu Rev Genet* **31**, 547-569.
- Glazebrook, J., Mitra, R., and Wang, L.** (2006). Signaling networks controlling, disease resistance responses in arabidopsis. *In Vitro Cell Dev-An* **42**, 11a-11a.
- Goehre, V., Spallek, T., Haeweker, H., Mersmann, S., Mentzel, T., Boller, T., de Torres, M., Mansfield, J.W., and Robatzek, S.** (2008). Plant Pattern-Recognition Receptor FLS2 Is Directed for Degradation by the Bacterial Ubiquitin Ligase *AvrPtoB*. *Current Biology* **18**, 1824-1832.
- Gough, C.** (2003). Rhizobium symbiosis: insight into Nod factor receptors. *Curr Biol* **13**, R973-975.

- Gust, A.A., Biswas, R., Lenz, H.D., Rauhut, T., Ranf, S., Kemmerling, B., Gotz, F., Glawischnig, E., Lee, J., Felix, G., and Nurnberger, T.** (2007). Bacteria-derived peptidoglycans constitute pathogen-associated molecular patterns triggering innate immunity in Arabidopsis. *J Biol Chem* **282**, 32338-32348.
- Hammond-Kosack, K.E., and Jones, J.D.G.** (1997). Plant disease resistance genes. *Annu Rev Plant Phys* **48**, 575-607.
- Hanaoka, H., Noda, T., Shirano, Y., Kato, T., Hayashi, H., Shibata, D., Tabata, S., and Ohsumi, Y.** (2002). Leaf senescence and starvation-induced chlorosis are accelerated by the disruption of an Arabidopsis autophagy gene. *Plant Physiol* **129**, 1181-1193.
- Harrison, C.J., and Langdale, J.A.** (2006). A step by step guide to phylogeny reconstruction. *Plant J* **45**, 561-572.
- Hayashi, F., Smith, K.D., Ozinsky, A., Hawn, T.R., Yi, E.C., Goodlett, D.R., Eng, J.K., Akira, S., Underhill, D.M., and Aderem, A.** (2001). The innate immune response to bacterial flagellin is mediated by Toll-like receptor 5. *Nature* **410**, 1099-1103.
- He, P., Shan, L., and Sheen, J.** (2007). Elicitation and suppression of microbe-associated molecular pattern-triggered immunity in plant-microbe interactions. *Cell Microbiol* **9**, 1385-1396.
- Heath, M.C.** (2000). Hypersensitive response-related death. *Plant Mol Biol* **44**, 321-334.
- Hedrick, S.A., Bell, J.N., Boller, T., and Lamb, C.J.** (1988). Chitinase Cdna Cloning and Messenger-Rna Induction by Fungal Elicitor, Wounding, and Infection. *Plant Physiol* **86**, 182-186.
- Hruz, T., Laule, O., Szaba, G., Wessendorp, F., and Bleuler, S.** (2008). Genevestigator V3: A Reference Expression Database for the Meta-Analysis of Transcriptomes. *Advances in Bioinformatics* **2008**, 5.
- Huang, Z.L., Yeakley, J.M., Garcia, E.W., Holdridge, J.D., Fan, J.B., and Whitham, S.A.** (2005). Salicylic acid-dependent expression of host genes in compatible Arabidopsis-virus interactions. *Plant Physiol* **137**, 1147-1159.
- Inohara, N., and Nunez, G.** (2003). NODs: intracellular proteins involved in inflammation and apoptosis. *Nat Rev Immunol* **3**, 371-382.
- Iuchi, S., Suzuki, H., Kim, Y.C., Iuchi, A., Kuromori, T., Ueguchi-Tanaka, M., Asami, T., Yamaguchi, I., Matsuoka, M., Kobayashi, M., and Nakajima, M.** (2007). Multiple loss-of-function of Arabidopsis gibberellin receptor AtGID1s completely shuts down a gibberellin signal. *Plant J* **50**, 958-966.
- Janjusevic, R., Abramovitch, R.B., Martin, G.B., and Stebbins, C.E.** (2006). A bacterial inhibitor of host programmed cell death defenses is an E3 ubiquitin ligase. *Science* **311**, 222-226.
- Jia, Y., McAdams, S.A., Bryan, G.T., Hershey, H.P., and Valent, B.** (2000). Direct interaction of resistance gene and avirulence gene products confers rice blast resistance. *Embo J* **19**, 4004-4014.
- Jones, J.D.G., and Dangl, J.L.** (2006). The plant immune system. *Nature* **444**, 323-329.
- Kaku, H., Nishizawa, Y., Ishii-Minami, N., Akimoto-Tomiyama, C., Dohmae, N., Takio, K., Minami, E., and Shibuya, N.** (2006). Plant cells recognize chitin fragments for defense signaling through a plasma membrane receptor. *Proc Natl Acad Sci U S A* **103**, 11086-11091.
- Katagiri, F., Thilmony, R., and He, S.Y.** (2002). The Arabidopsis Thaliana-Pseudomonas Syringae Interaction. *The Arabidopsis Book*, 1-35.

- Kearsey, M.J., and Pooni, H.S.** (1996). The genetical analysis of quantitative traits. (Chapman and Hall).
- Kim, M.G., da Cunha, L., McFall, A.J., Belkhadir, Y., DebRoy, S., Dangl, J.L., and Mackey, D.** (2005a). Two *Pseudomonas syringae* type III effectors inhibit RIN4-regulated basal defense in *Arabidopsis*. *Cell* **121**, 749-759.
- Kim, M.G., da Cunha, L., McFall, A.J., Belkhadir, Y., DebRoy, S., Dangl, J.L., and Mackey, D.** (2005b). Two *Pseudomonas syringae* type III effectors inhibit RIM-regulated basal defense in *Arabidopsis*. *Cell* **121**, 749-759.
- Knogge, W., and Scheel, D.** (2006). LysM receptors recognize friend and foe. *Proc Natl Acad Sci U S A* **103**, 10829-10830.
- Kobe, B., and Kajava, A.V.** (2001). The leucine-rich repeat as a protein recognition motif. *Curr Opin Struc Biol* **11**, 725-732.
- Koornneef, A., and Pieterse, C.M.** (2008). Cross talk in defense signaling. *Plant Physiol* **146**, 839-844.
- Koornneef, M., Alonso-Blanco, C., and Vreugdenhil, D.** (2004). Naturally occurring genetic variation in *Arabidopsis thaliana*. *Annual Review of Plant Biology* **55**, 141-172.
- Koornneef, M., Alonso-Blanco, C., Peeters, A.J.M., and Soppe, W.** (1998). Genetic control of flowering time in *arabidopsis*. *Annu Rev Plant Phys* **49**, 345-370.
- Kover, P.X., and Schaal, B.A.** (2002). Genetic variation for disease resistance and tolerance among *Arabidopsis thaliana* accessions. *P Natl Acad Sci USA* **99**, 11270-11274.
- Kover, P.X., Wolf, J.B., Kunkel, B.N., and Cheverud, J.M.** (2005). Genetic architecture of *Arabidopsis thaliana* response to infection by *Pseudomonas syringae*. *Heredity* **94**, 507-517.
- Kowalski, S.P., Lan, T.H., Feldmann, K.A., and Paterson, A.H.** (1994). Qtl Mapping of Naturally-Occurring Variation in Flowering Time of *Arabidopsis-Thaliana*. *Mol Gen Genet* **245**, 548-555.
- Krysan, P.J., Young, J.C., and Sussman, M.R.** (1999). T-DNA as an insertional mutagen in *Arabidopsis*. *Plant Cell* **11**, 2283-2290.
- Kunze, G., Zipfel, C., Robatzek, S., Niehaus, K., Boller, T., and Felix, G.** (2004). The N terminus of bacterial elongation factor Tu elicits innate immunity in *Arabidopsis* plants. *Plant Cell* **16**, 3496-3507.
- Kwon, C.** (2010). Plant Defense Responses Coming To Shape. *The Plant Pathology Journal* Volume **26**, No **2**, pp. **115-120**.
- Labanna, K.S., and Banga, S.S.** (1993). Breeding oilseed brassicas. *Monographs on Theoretical and applied genetics* **19**.
- Lacombe, S., Rougon-Cardoso, A., Sherwood, E., Peeters, N., Dahlbeck, D., van Esse, H.P., Smoker, M., Rallapalli, G., Thomma, B.P., Staskawicz, B., Jones, J.D., and Zipfel, C.** (2010). Interfamily transfer of a plant pattern-recognition receptor confers broad-spectrum bacterial resistance. *Nat Biotechnol* **28**, 365-369.
- Lamb, C., and Dixon, R.A.** (1997). The oxidative burst in plant disease resistance. *Annu Rev Plant Phys* **48**, 251-275.
- Laurie, D.A.** (1997). Comparative genetics of flowering time. *Plant Mol Biol* **35**, 167-177.
- Lehmann, E.L.** (1975). *Nonparametrics: Statistical methods based on ranks*. (San Francisco: Holden-Day).

- Lewis, J.D., Guttman, D.S., and Desveaux, D.** (2009). The targeting of plant cellular systems by injected type III effector proteins. *Semin Cell Dev Biol* **20**, 1055-1063.
- Li, H., Deng, Y., Wu, T., Subramanian, S., and Yu, O.** (2010). Misexpression of miR482, miR1512, and miR1515 increases soybean nodulation. *Plant Physiol* **153**, 1759-1770.
- Lin, N.C., and Martin, G.B.** (2005). An *avrPto/avrPtoB* mutant of *Pseudomonas syringae* pv. tomato DC3000 does not elicit Pto-mediated resistance and is less virulent on tomato. *Mol Plant Microbe Interact* **18**, 43-51.
- Lipka, U., Fuchs, R., Kuhns, C., Petutschnig, E., and Lipka, V.** (2010). Live and let die--*Arabidopsis* nonhost resistance to powdery mildews. *Eur J Cell Biol* **89**, 194-199.
- Lipka, V., Kwon, C., and Panstruga, R.** (2007). SNARE-Ware: The Role of SNARE-Domain Proteins in Plant Biology. *Annu Rev Cell Dev Biol*.
- Liu, B.H.** (1998). *Statistical Genomics: Linkage, Mapping and QTL Analysis*. (CRC Press).
- Lohmann, G.V., Shimoda, Y., Nielsen, M.W., Jorgensen, F.G., Grossmann, C., Sandal, N., Sorensen, K., Thirup, S., Madsen, L.H., Tabata, S., Sato, S., Stougaard, J., and Radutoiu, S.** (2010). Evolution and regulation of the *Lotus japonicus* LysM receptor gene family. *Mol Plant Microbe Interact* **23**, 510-521.
- Lu, M., Tang, X.Y., and Zhou, J.M.** (2001). *Arabidopsis* NHO1 is required for general resistance against *Pseudomonas* bacteria. *Plant Cell* **13**, 437-447.
- Mackey, D., Belkhadir, Y., Alonso, J.M., Ecker, J.R., and Dangl, J.L.** (2003). *Arabidopsis* RIN4 is a target of the type III virulence effector AvrRpt2 and modulates RPS2-mediated resistance. *Cell* **112**, 379-389.
- Madsen, E.B., Madsen, L.H., Radutoiu, S., Olbryt, M., Rakwalska, M., Szczyglowski, K., Sato, S., Kaneko, T., Tabata, S., Sandal, N., and Stougaard, J.** (2003). A receptor kinase gene of the LysM type is involved in legume perception of rhizobial signals. *Nature* **425**, 637-640.
- Maloof, J.N., Borevitz, J.O., Dabi, T., Lutes, J., Nehring, R.B., Redfern, J.L., Trainer, G.T., Wilson, J.M., Asami, T., Berry, C.C., Weigel, D., and Chory, J.** (2001). Natural variation in light sensitivity of *Arabidopsis*. *Nat Genet* **29**, 441-446.
- Mauricio, R., Stahl, E.A., Korves, T., Tian, D.C., Kreitman, M., and Bergelson, J.** (2003). Natural selection for polymorphism in the disease resistance gene *Rps2* of *Arabidopsis thaliana*. *Genetics* **163**, 735-746.
- McDonald, B.A., and Linde, C.** (2002). The population genetics of plant pathogens and breeding strategies for durable resistance. *Euphytica* **124**, 163-180.
- Meyers, B.C., Dickerman, A.W., Michelmore, R.W., Sivaramakrishnan, S., Sobral, B.W., and Young, N.D.** (1999). Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide-binding superfamily. *Plant J* **20**, 317-332.
- Mishina, T.E., and Zeier, J.** (2006). The *Arabidopsis* flavin-dependent monooxygenase FMO1 is an essential component of biologically induced systemic acquired resistance. *Plant Physiol* **141**, 1666-1675.
- Mitchell-Olds, T., and Schmitt, J.** (2006). Genetic mechanisms and evolutionary significance of natural variation in *Arabidopsis*. *Nature* **441**, 947-952.
- Miya, A., Albert, P., Shinya, T., Desaki, Y., Ichimura, K., Shirasu, K., Narusaka, Y., Kawakami, N., Kaku, H., and Shibuya, N.** (2007). CERK1, a LysM

- receptor kinase, is essential for chitin elicitor signaling in Arabidopsis. *Proc Natl Acad Sci U S A* **104**, 19613-19618.
- Moran, P.J., and Thompson, G.A.** (2001). Molecular responses to aphid feeding in Arabidopsis in relation to plant defense pathways. *Plant Physiol* **125**, 1074-1085.
- Morel, J.B., and Dangl, J.L.** (1997). The hypersensitive response and the induction of cell death in plants. *Cell Death Differ* **4**, 671-683.
- Nandi, A., Kachroo, P., Fukushige, H., Hildebrand, D.F., Klessig, D.F., and Shah, J.** (2003). Ethylene and jasmonic acid signaling affect the NPR1-independent expression of defense genes without impacting resistance to *Pseudomonas syringae* and *Peronospora parasitica* in the Arabidopsis *ssi1* mutant. *Mol Plant Microbe In* **16**, 588-599.
- Navarro, L., Zipfel, C., Rowland, O., Keller, I., Robatzek, S., Boller, T., and Jones, J.D.G.** (2004). The transcriptional innate immune response to flg22. interplay and overlap with Avr gene-dependent defense responses and bacterial pathogenesis. *Plant Physiol* **135**, 1113-1128.
- Nekrasov, V., Li, J., Batoux, M., Roux, M., Chu, Z.H., Lacombe, S., Rougon, A., Bittel, P., Kiss-Papp, M., Chinchilla, D., van Esse, H.P., Jorda, L., Schwessinger, B., Nicaise, V., Thomma, B.P., Molina, A., Jones, J.D., and Zipfel, C.** (2009). Control of the pattern-recognition receptor EFR by an ER protein complex in plant immunity. *Embo J* **28**, 3428-3438.
- Nicaise, V., Roux, M., and Zipfel, C.** (2009). Recent advances in PAMP-triggered immunity against bacteria: pattern recognition receptors watch over and raise the alarm. *Plant Physiol* **150**, 1638-1647.
- Nimchuk, Z., Rohmer, L., Chang, J.H., and Dangl, J.L.** (2001). Knowing the dancer from the dance: R-gene products and their interactions with other proteins from host and pathogen. *Curr Opin Plant Biol* **4**, 288-294.
- Nurnberger, T., and Kemmerling, B.** (2006). Receptor protein kinases--pattern recognition receptors in plant immunity. *Trends Plant Sci* **11**, 519-522.
- Oh, H.S., and Collmer, A.** (2005). Basal resistance against bacteria in *Nicotiana benthamiana* leaves is accompanied by reduced vascular staining and suppressed by multiple *Pseudomonas syringae* type III secretion system effector proteins. *Plant J* **44**, 348-359.
- Osbourn, A.E.** (1996). Preformed Antimicrobial Compounds and Plant Defense against Fungal Attack. *Plant Cell* **8**, 1821-1831.
- Osbourn, A.E.** (1999). Antimicrobial phytoprotectants and fungal pathogens: a commentary. *Fungal Genet Biol* **26**, 163-168.
- Ossowski, S., Schwab, R., and Weigel, D.** (2008). Gene silencing in plants using artificial microRNAs and other small RNAs. *Plant J* **53**, 674-690.
- Park, Y.S., Jeon, M.H., Lee, S.H., Moon, J.S., Cha, J.S., Kim, H.Y., and Cho, T.J.** (2005). Activation of defense responses in Chinese cabbage by a nonhost pathogen, *Pseudomonas syringae* pv. tomato. *J Biochem Mol Biol* **38**, 748-754.
- Parker, J.E., Coleman, M.J., Szabo, V., Frost, L.N., Schmidt, R., vanderBiezen, E.A., Moores, T., Dean, C., Daniels, M.J., and Jones, J.D.G.** (1997). The Arabidopsis downy mildew resistance gene RPP5 shares similarity to the toll and interleukin-1 receptors with N and L6. *Plant Cell* **9**, 879-894.
- Penninckx, I.A.M.A., Thomma, B.P.H.J., Buchala, A., Metraux, J.P., and Broekaert, W.F.** (1998). Concomitant activation of jasmonate and ethylene

- response pathways is required for induction of a plant defensin gene in *Arabidopsis*. *Plant Cell* **10**, 2103-2113.
- Pieterse, C.M.J., and van Loon, L.C.** (1999). Salicylic acid-independent plant defence pathways. *Trends Plant Sci* **4**, 52-58.
- Pieterse, C.M.J., Ton, J., and Van Loon, L.C.** (2001). Cross-talk between plant defence signalling pathways: boost or burden? *AgBiotechNet* **3**.
- Plude, J.L., Parker, D.L., Schommer, O.J., Timmerman, R.J., Hagstrom, S.A., Joers, J.M., and Hnasko, R.** (1991). Chemical Characterization of Polysaccharide from the Slime Layer of the Cyanobacterium *Microcystis-Flos-Aquae-C3-40*. *Appl Environ Microb* **57**, 1696-1700.
- Qiu, J.L., Fiil, B.K., Petersen, K., Nielsen, H.B., Botanga, C.J., Thorgrimsen, S., Palma, K., Suarez-Rodriguez, M.C., Sandbech-Clausen, S., Lichota, J., Brodersen, P., Grasser, K.D., Mattsson, O., Glazebrook, J., Mundy, J., and Petersen, M.** (2008). *Arabidopsis* MAP kinase 4 regulates gene expression through transcription factor release in the nucleus. *Embo J* **27**, 2214-2221.
- Radutoiu, S., Madsen, L.H., Madsen, E.B., Jurkiewicz, A., Fukai, E., Quistgaard, E.M.H., Albrechtsen, A.S., James, E.K., Thirup, S., and Stougaard, J.** (2007). LysM domains mediate lipochitin-oligosaccharide recognition and *Nfr* genes extend the symbiotic host range. *Embo Journal* **26**, 3923-3935.
- Radutoiu, S., Madsen, L.H., Madsen, E.B., Felle, H.H., Umehara, Y., Gronlund, M., Sato, S., Nakamura, Y., Tabata, S., Sandal, N., and Stougaard, J.** (2003). Plant recognition of symbiotic bacteria requires two LysM receptor-like kinases. *Nature* **425**, 585-592.
- Rentel, M.C., Lecourieux, D., Ouaked, F., Usher, S.L., Petersen, L., Okamoto, H., Knight, H., Peck, S.C., Grierson, C.S., Hirt, H., and Knight, M.R.** (2004). OXI1 kinase is necessary for oxidative burst-mediated signalling in *Arabidopsis*. *Nature* **427**, 858-861.
- Ryals, J.A., Neuenschwander, U.H., Willits, M.G., Molina, A., Steiner, H.-Y., and Hunt, M.D.** (1996). Systemic Acquired Resistance. *The Plant Cell Vol. 8*, 1809-1819.
- Schenk, P.M., Kazan, K., Rusu, A.G., Manners, J.M., and Maclean, D.J.** (2005). The SEN1 gene of *Arabidopsis* is regulated by signals that link plant defence responses and senescence. *Plant Physiol Bioch* **43**, 997-1005.
- Schenk, P.M., Kazan, K., Wilson, I., Anderson, J.P., Richmond, T., Somerville, S.C., and Manners, J.M.** (2000). Coordinated plant defense responses in *Arabidopsis* revealed by microarray analysis. *Proc Natl Acad Sci U S A* **97**, 11655-11660.
- Schmid, K.J., Sorensen, T.R., Stracke, R., Torjek, O., Altmann, T., Mitchell-Olds, T., and Weisshaar, B.** (2003a). Large-scale identification and analysis of genome-wide single-nucleotide polymorphisms for mapping in *Arabidopsis thaliana*. *Genome Res* **13**, 1250-1257.
- Schmid, M., Uhlenhaut, N.H., Godard, F., Demar, M., Bressan, R., Weigel, D., and Lohmann, J.U.** (2003b). Dissection of floral induction pathways using global expression analysis. *Development* **130**, 6001-6012.
- Schwab, R., Ossowski, S., Riester, M., Warthmann, N., and Weigel, D.** (2006). Highly specific gene silencing by artificial microRNAs in *Arabidopsis*. *Plant Cell* **18**, 1121-1133.
- Schwessinger, B., and Zipfel, C.** (2008). News from the frontline: recent insights into PAMP-triggered immunity in plants. *Curr Opin Plant Biol* **11**, 389-395.

- Shan, L.B., He, P., and Sheen, J.** (2007). Endless hide-and-seek: Dynamic co-evolution in plant-bacterium warfare. *J Integr Plant Biol* **49**, 105-111.
- Smit, P., Limpens, E., Geurts, R., Fedorova, E., Dolgikh, E., Gough, C., and Bisseling, T.** (2007). Medicago LYK3, an entry receptor in rhizobial nodulation factor signaling. *Plant Physiol* **145**, 183-191.
- Spoel, S.H., Mou, Z., Tada, Y., Spivey, N.W., Genschik, P., and Dong, X.** (2009). Proteasome-mediated turnover of the transcription coactivator NPR1 plays dual roles in regulating plant immunity. *Cell* **137**, 860-872.
- Sprague, S.J., Balesdent, M.H., Brun, H., Hayden, H.L., Marcroft, S.J., Pinochet, X., Rouxel, T., and Howlett, B.J.** (2006). Major gene resistance in *Brassica napus* (oilseed rape) is overcome by changes in virulence of populations of *Leptosphaeria maculans* in France and Australia. *Eur J Plant Pathol* **114**, 33-40.
- Staswick, P.E., Yuen, G.Y., and Lehman, C.C.** (1998). Jasmonate signaling mutants of *Arabidopsis* are susceptible to the soil fungus *Pythium irregulare*. *Plant J* **15**, 747-754.
- Stintzi, A., Heitz, T., Prasad, V., Wiedemannmerdinoglu, S., Kauffmann, S., Geoffroy, P., Legrand, M., and Fritig, B.** (1993). Plant Pathogenesis-Related Proteins and Their Role in Defense against Pathogens. *Biochimie* **75**, 687-706.
- Suarez-Rodriguez, M.C., Adams-Phillips, L., Liu, Y., Wang, H., Su, S.H., Jester, P.J., Zhang, S., Bent, A.F., and Krysan, P.J.** (2007). MEKK1 is required for flg22-induced MPK4 activation in *Arabidopsis* plants. *Plant Physiol* **143**, 661-669.
- Tada, Y., Spoel, S.H., Pajerowska-Mukhtar, K., Mou, Z., Song, J., Wang, C., Zuo, J., and Dong, X.** (2008). Plant immunity requires conformational changes [corrected] of NPR1 via S-nitrosylation and thioredoxins. *Science* **321**, 952-956.
- Taiz, L., and Zeiger, E.** (2002). *Plant Physiology*. (Sinauer Associates, Inc.).
- Tang, X., Xie, M., Kim, Y.J., Zhou, J., Klessig, D.F., and Martin, G.B.** (1999). Overexpression of Pto activates defense responses and confers broad resistance. *Plant Cell* **11**, 15-29.
- Tax, F.E., and Vernon, D.M.** (2001). T-DNA-associated duplication/translocations in *Arabidopsis*. Implications for mutant analysis and functional genomics. *Plant Physiol* **126**, 1527-1538.
- Thaler, J.S., Karban, R., Ullman, D.E., Boege, K., and Bostock, R.M.** (2002). Cross-talk between jasmonate and salicylate plant defense pathways: effects on several plant parasites. *Oecologia* **131**, 227-235.
- Thomma, B.P., Eggermont, K., Penninckx, I.A., Mauch-Mani, B., Vogelsang, R., Cammue, B.P., and Broekaert, W.F.** (1998). Separate jasmonate-dependent and salicylate-dependent defense-response pathways in *Arabidopsis* are essential for resistance to distinct microbial pathogens. *Proc Natl Acad Sci U S A* **95**, 15107-15111.
- Thordal-Christensen, H., Zhang, Z.G., Wei, Y.D., and Collinge, D.B.** (1997). Subcellular localization of H₂O₂ in plants. H₂O₂ accumulation in papillae and hypersensitive response during the barley-powdery mildew interaction. *Plant J* **11**, 1187-1194.
- Ton, J., Davison, S., Van Loon, L.C., and Pieterse, C.M.J.** (2001). Heritability of rhizobacteria-mediated induced systemic resistance and basal resistance in *Arabidopsis*. *Eur J Plant Pathol* **107**, 63-68.

- Tornero, P., Merritt, P., Sadanandom, A., Shirasu, K., Innes, R.W., and Dangl, J.L.** (2002). RAR1 and NDR1 contribute quantitatively to disease resistance in Arabidopsis, and their relative contributions are dependent on the R gene assayed. *Plant Cell* **14**, 1005-1015.
- Torres, M.A., Dangl, J.L., and Jones, J.D.G.** (2002). Arabidopsis gp91(phox) homologues AtrbohD and AtrbohF are required for accumulation of reactive oxygen intermediates in the plant defense response. *P Natl Acad Sci USA* **99**, 517-522.
- Tsuda, K., Sato, M., Glazebrook, J., Cohen, J.D., and Katagiri, F.** (2008). Interplay between MAMP-triggered and SA-mediated defense responses. *Plant J* **53**, 763-775.
- Tsuda, K., Sato, M., Stoddard, T., Glazebrook, J., and Katagiri, F.** (2009). Network properties of robust immunity in plants. *PLoS Genet* **5**, e1000772.
- Tyler, L., Thomas, S.G., Hu, J.H., Dill, A., Alonso, J.M., Ecker, J.R., and Sun, T.P.** (2004). DELLA proteins and gibberellin-regulated seed germination and floral development in Arabidopsis. *Plant Physiol* **135**, 1008-1019.
- Ulker, B., Peiter, E., Dixon, D.P., Moffat, C., Capper, R., Bouche, N., Edwards, R., Sanders, D., Knight, H., and Knight, M.R.** (2008). Getting the most out of publicly available T-DNA insertion lines. *Plant J*.
- Van der Plank, J.E.** (1963). *Plant diseases: Epidemics and control*. Academic Press, 349.
- Van Ooijen, J., and Maliepaard, C.** (1996). MapQTL (tm) version 3.0: Software for the calculation of QTL positions on genetic maps (CPRO-DLO).
- Van Ooijen, J.W.** (1999). LOD significance thresholds for QTL analysis in experimental populations of diploid species. *Heredity* **83 (Pt 5)**, 613-624.
- Van Wees, S.C., De Swart, E.A., Van Pelt, J.A., Van Loon, L.C., and Pieterse, C.M.** (2000). Enhancement of induced disease resistance by simultaneous activation of salicylate- and jasmonate-dependent defense pathways in Arabidopsis thaliana. *Proc Natl Acad Sci U S A* **97**, 8711-8716.
- VanderVliet, A., Eiserich, J.P., Halliwell, B., and Cross, C.E.** (1997). Formation of reactive nitrogen species during peroxidase-catalyzed oxidation of nitrite - A potential additional, mechanism of nitric oxide-dependent toxicity. *J Biol Chem* **272**, 7617-7625.
- Varga, G.J., and Szegedi, E.** (2007). Basal resistance: an alternative defence strategy against crown gall disease. Research Institute for Viticulture and Enology of Ministry of Agricultural and Rural Development, Kecskemét, POBox 25, H-6001, Hungary varga.g@szbkik.hu, phone: +36/76/494888.
- Veronese, P., Ruiz, M.T., Coca, M.A., Hernandez-Lopez, A., Lee, H., Ibeas, J.I., Damsz, B., Pardo, J.M., Hasegawa, P.M., Bressan, R.A., and Narasimhan, M.L.** (2003). In defense against pathogens. Both plant sentinels and foot soldiers need to know the enemy. *Plant Physiol* **131**, 1580-1590.
- Wan, J.R., Zhang, X.C., Neece, D., Ramonell, K.M., Clough, S., Kim, S.Y., Stacey, M.G., and Stacey, G.** (2008). A LysM receptor-like kinase plays a critical role in chitin signaling and fungal resistance in Arabidopsis. *Plant Cell* **20**, 471-481.
- Wang, D., Amornsiripanitch, N., and Dong, X.** (2006). A genomic approach to identify regulatory nodes in the transcriptional network of systemic acquired resistance in plants. *PLoS Pathog* **2**, e123.
- Weigel, D., and Nordborg, M.** (2005). Natural variation in arabidopsis. How do we find the causal genes? *Plant Physiol* **138**, 567-568.

- Wildermuth, M.C., Dewdney, J., Wu, G., and Ausubel, F.M.** (2001). Isochorismate synthase is required to synthesize salicylic acid for plant defence. *Nature* **414**, 562-565.
- Winson, M.K., Swift, S., Hill, P.J., Sims, C.M., Griesmayr, G., Bycroft, B.W., Williams, P., and Stewart, G.S.A.B.** (1998). Engineering the luxCDABE genes from *Photobacterium luminescens* to provide a bioluminescent reporter for constitutive and promoter probe plasmids and mini-Tn5 constructs. *FEMS Microbiol Lett* **163**, 193-202.
- Xiao, S.Y., Ellwood, S., Calis, O., Patrick, E., Li, T.X., Coleman, M., and Turner, J.G.** (2001). Broad-spectrum mildew resistance in *Arabidopsis thaliana* mediated by RPW8. *Science* **291**, 118-120.
- Xiao, S.Y., Calis, O., Patrick, E., Zhang, G.G., Charoenwattana, P., Muskett, P., Parker, J.E., and Turner, J.G.** (2005). The atypical resistance gene, RPW8, recruits components of basal defence for powdery mildew resistance in *Arabidopsis*. *Plant J* **42**, 95-110.
- Xu, X.P., Chen, C.H., Fan, B.F., and Chen, Z.X.** (2006). Physical and functional interactions between pathogen-induced *Arabidopsis* WRKY18, WRKY40, and WRKY60 transcription factors. *Plant Cell* **18**, 1310-1326.
- Yang, C.W., Gonzalez-Lamothe, R., Ewan, R.A., Rowland, O., Yoshioka, H., Shenton, M., Ye, H., O'Donnell, E., Jones, J.D., and Sadanandom, A.** (2006). The E3 ubiquitin ligase activity of *Arabidopsis* PLANT U-BOX17 and its functional tobacco homolog ACRE276 are required for cell death and defense. *Plant Cell* **18**, 1084-1098.
- Yoshida, H., Kinoshita, K., and Ashida, M.** (1996). Purification of a peptidoglycan recognition protein from hemolymph of the silkworm, *Bombyx mori*. *J Biol Chem* **271**, 13854-13860.
- Young, J.C., Krysan, P.J., and Sussman, M.R.** (2001). Efficient screening of *Arabidopsis* T-DNA insertion lines using degenerate primers. *Plant Physiol* **125**, 513-518.
- Young, N.D.** (1996). QTL mapping and quantitative disease resistance in plants. *Annu Rev Phytopathol* **34**, 479-501.
- Yu, T.S., Zeeman, S.C., Thorneycroft, D., Fulton, D.C., Dunstan, H., Lue, W.L., Hegemann, B., Tung, S.Y., Umemoto, T., Chapple, A., Tsai, D.L., Wang, S.M., Smith, A.M., Chen, J., and Smith, S.M.** (2005). α -Amylase is not required for breakdown of transitory starch in *Arabidopsis* leaves. *J Biol Chem* **280**, 9773-9779.
- Zhang, X.C., Wu, X., Findley, S., Wan, J., Libault, M., Nguyen, H.T., Cannon, S.B., and Stacey, G.** (2007). Molecular evolution of lysin motif-type receptor-like kinases in plants. *Plant Physiol* **144**, 623-636.
- Zhou, N., Tootle, T.L., Tsui, F., Klessig, D.F., and Glazebrook, J.** (1998). PAD4 functions upstream from salicylic acid to control defense responses in *Arabidopsis*. *Plant Cell* **10**, 1021-1030.
- Zhu, X., and Galili, G.** (2003). Increased lysine synthesis coupled with a knockout of its catabolism synergistically boosts lysine content and also transregulates the metabolism of other amino acids in *Arabidopsis* seeds. *Plant Cell* **15**, 845-853.
- Zhu, X., Santat, L.A., Chang, M.S., Liu, J., Zavzavadjian, J.R., Wall, E.A., Kivork, C., Simon, M.I., and Fraser, I.D.** (2007). A versatile approach to multiple gene RNA interference using microRNA-based short hairpin RNAs. *BMC Mol Biol* **8**, 98.

- Zimmermann, P., Hirsch-Hoffmann, M., Hennig, L., and Gruissem, W.** (2004). GENEVESTIGATOR. Arabidopsis microarray database and analysis toolbox. *Plant Physiol* **136**, 2621-2632.
- Zipfel, C.** (2008). Pattern-recognition receptors in plant innate immunity. *Curr Opin Immunol* **20**, 10-16.
- Zipfel, C.** (2009). Early molecular events in PAMP-triggered immunity. *Curr Opin Plant Biol* **12**, 414-420.
- Zipfel, C., and Felix, G.** (2005). Plants and animals: a different taste for microbes? *Curr Opin Plant Biol* **8**, 353-360.
- Zipfel, C., and Rathjen, J.P.** (2008). Plant immunity: AvrPto targets the frontline. *Curr Biol* **18**, R218-220.
- Zipfel, C., Robatzek, S., Navarro, L., Oakeley, E.J., Jones, J.D.G., Felix, G., and Boller, T.** (2004). Bacterial disease resistance in Arabidopsis through flagellin perception. *Nature* **428**, 764-767.
- Zipfel, C., Kunze, G., Chinchilla, D., Caniard, A., Jones, J.D., Boller, T., and Felix, G.** (2006). Perception of the bacterial PAMP EF-Tu by the receptor EFR restricts Agrobacterium-mediated transformation. *Cell* **125**, 749-760.
- Zwieniecki, M.A., Orians, C.M., Melcher, P.J., and Holbrook, N.M.** (2003). Ionic control of the lateral exchange of water between vascular bundles in tomato. *J Exp Bot* **54**, 1399-1405.

Appendices

1.1 PCR primers